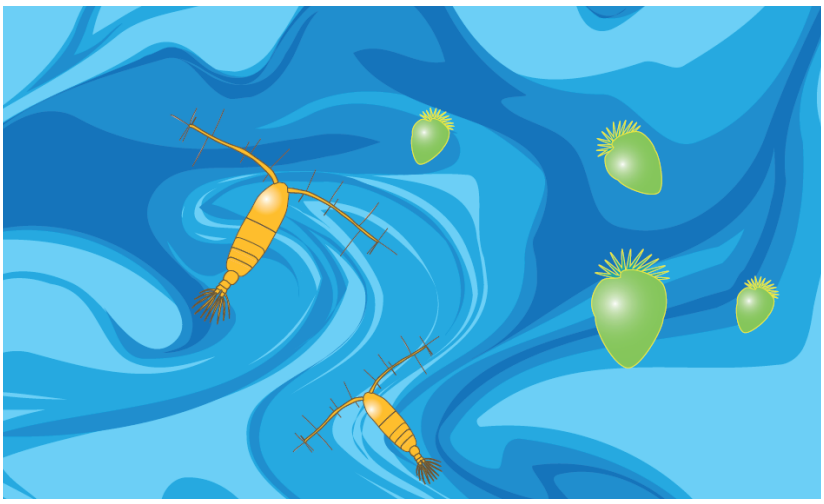
Small, wet & rational

Individual based zooplankton ecology

Visser, Andre

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record



Small, wet & rational

Individual based

zooplankton ecology

André W. Visser

May 8, 2011

Resumé

Small, wet and rational: Individual-based zooplankton ecology.

Marine ecosystems are a weave of many strands, the most fundamental of which is the life trajectory of an individual planktonic organism. The topology of such a trajectory is a function of the environmental factors an organism encounters (both biotic and abiotic), its physical abilities, and the behaviour it adopts. These latter attributes (abilities and behavioural algorithms) are a consequence of natural selection, and should therefore reflect the evolutionary tradeoffs that shaped them. This monograph couples together principles from hydrodynamics, oceanography, evolutionary and behavioural ecology to provide a mechanistic rationale for the adaptive behaviour of zooplankton, and the consequences of fitness-seeking behaviour on the dynamics of marine communities. This development starts with the observation that all important life processes in the plankton (growth, reproduction, mortality) depend on encounter rates. Encounter rates represent not only a potential transfer biomass, but more importantly, an exchange of information - information that when processed, trigger behavioural responses that are attuned to an organism's evolutionary self-interest. A powerful concept here is a mechanistic link between behaviour and Darwinian fitness (*e.g.* the trade-off between benefit, cost and risk of foraging). Relatively simple arguments along these lines give insights as to why plankters behave the way they do: how fast they swim, their optimal path geometry, vertical positioning and time allocation in foraging. It also resolves some aspects of population dynamics – the paradoxes of omnivory and enrichment for instance – in that adaptive (*i.e.* fitness-seeking) behaviour has a stabilizing effect on community dynamics.

Contents

Resumé

Small, wet and rational: Individual-based zooplankton ecology.

Marine ecosystems are a weave of many strands, the most fundamental of which is the life trajectory of an individual planktonic organism. The topology of such a trajectory is a function of the environmental factors an organism encounters (both biotic and abiotic), its physical abilities, and the behaviour it adopts. These latter attributes (abilities and behavioural algorithms) are a consequence of natural selection, and should therefore reflect the evolutionary tradeoffs that shaped them. This monograph couples together principles from hydrodynamics, oceanography, evolutionary and behavioural ecology to provide a mechanistic rationale for the adaptive behaviour of zooplankton, and the consequences of fitness-seeking behaviour on the dynamics of marine communities. This development starts with the observation that all important life processes in the plankton (growth, reproduction, mortality) depend on encounter rates. Encounter rates represent not only a potential transfer biomass, but more importantly, an exchange of information - information that when processed, trigger behavioural responses that are attuned to an organism's evolutionary self-interest. A powerful concept here is a mechanistic link between behaviour and Darwinian fitness (*e.g.* the trade-off between benefit, cost and risk of foraging). Relatively simple arguments along these lines give insights as to why plankters behave the way they do: how fast they swim, their optimal path geometry, vertical positioning and time allocation in foraging. It also resolves some aspects of population dynamics – the paradoxes of omnivory and enrichment for instance – in that adaptive (*i.e.* fitness-seeking) behaviour has a stabilizing effect on community dynamics.

Lille, våd og rationel: Individbaseret zooplanktonøkologi.

Marine økosystemer er et tæppe vævet af mange tråde, hvoraf den mest grundlæggende er livsforløbet hos den enkelte planktonorganisme. Et sådant livsforløb er en funktion af de miljømæssige faktorer, som en organisme møder (både biotiske og abiotiske), dens fysiske evner og dens adfærd. De sidstnævnte egenskaber (evner og adfærd) er et resultat af naturlig udvælgelse og skulle derfor afspejle de udviklingsmæssige hensyn som formede dem. Denne monografi udnytter principperne fra hydrodynamik, oceanografi, evolutionær og adfærdsmæssig økologi til at give en mekanistisk logisk begrundelse for zooplanktons tilpasningsadfærd og følgerne af fitness optimering på dynamikken i marine samfund. Alle vigtige livsprocesser i plankton (vækst, reproduktion, dødelighed) afhænger af kontaktrater, som ikke blot udgør en mulig overførsel af biomasse, men endnu vigtigere: en udveksling af informationer, som - når de bliver behandlet - udløser adfærdsmæssige reaktioner, der er afstemt efter en organismes udviklingsmæssige fordele. Et vigtigt begreb her er, at der er en mekanistisk forbindelse mellem adfærd og darwinistisk tilpasningsevne (f.eks. afvejning af fordele, ulemper og risici ved fødeindtagelse). Relativt enkle argumenter af denne type giver øget viden om, hvorfor plankton opfører sig som de gør: hvor hurtigt de svømmer, deres optimale svømmemønster, vertikal positionering i vandsøjlen og tidsforbrug til fouragering. Det løser derudover nogle af populationsdynamikkens paradokser – f.eks. "paradox of omnivory" og "paradox of enrichment" – idet dynamisk adfærdsoptimering har en stabiliserende effekt på samfundets dynamik.

1 Introduction

Small, Wet and Rational:

The title of this work reflects 3 important aspects of plankton; they are small, at least in comparison to us, they inhabit a fluid environment, and while they may not be "thinking" creatures, they are endowed by natural selection with behavioural algorithms that are open to rational interpretation. Small and wet also underlines the idea of the individual plankter. It is at the individual scale that the immediate impact of the fluid environment is manifest. How do planktonic organisms move, find resources, detect and capture prey, find mates, escape from predators? These are fundamental life processes that can only be fully understood from the perspective of the individual and the small scale fluid dynamics of its immediate environment. It is also the individual scale at which natural selection is active – attributes that lend individuals greater success in meeting the challenges of survival and reproduction will have a greater expression in succeeding generations. This is true also for behavioural algorithms. It is my contention that by bringing behavioural and evolutionary ecology together at the individual level, many of aspects of planktonic interactions can be given a mechanistic understanding. That is, if we can deduce the rules that shape behaviour, then we have extra information through which behavioural changes can be predicted. Zooplankton have rational behaviour, that has meaning only for the individual, but has implications across communities, populations and ecosystems.

This is the theme I wish to promote with this work. In order to do so, I will delve into the physical realm (small and wet) within which plankton live. The first 3 chapters deal specifically with this. The first is a treatment of encounter rates – the rate at which an individual plankter contacts other stuff – either other organisms, detrital particles or patches of resource. Encounter rates are the fundamental currency of all planktonic interactions and are determined by

- how densely packed targets are in space
- how fast targets move relative to the searcher
- how far a target can be detected by a searcher

where "searchers" and "targets" are only convenient labels for actors that can take on any number roles in specific planktonic interactions: a predator "searches" for a prey "target", but equally, a prey can be alert for and thus "search" for a predator "target". Encounter rates also depend on motility patterns. To what degree the swimming track of an organism is straight-line or zig-zag impacts their encounter rate with predators and prey; an aspect that can be exploited in their foraging strategy.

The following 2 chapters deal with the fluid nature of the environment that plankton inhabit. The physics of this environment as experienced by plankton is alien to our sense of the same environment. For plankton, being small, water is sticky. This is a curious feature of small scale fluid dynamics (at least to our experience) that is usually alluded to. But so too are buoyancy, pressure and density. Plankton are cradled in their watery environments, and whatever forces make the water move, act on the plankton as well. Most of the life drama of plankton is played out in the viscous realm of low Reynolds numbers. While foreign to us, zooplankton are well adapted to to their world, in their means of locomotion, but also in harvesting information; the hydromechanical signals and chemical trails that other planktonic organisms invariably leave behind. Such information, how it is produced by a "target", attenuated by the environment, and sensed and interpreted by the "searcher" are fundamental to the interaction between plankters.

The following chapter on turbulence recognizes that while individual plankters may experience mostly low Reynolds number, their environment certainly does not. Mechanical energy cascades down from storms, winds, waves and tides to be dissipated as heat at smaller scales – as it turns out, at a size about that of a typical zooplankton. Turbulent motion impacts the marine environment across all scales – it stirs and mixes material and biota, bringing planktonic organisms into contact with each other and shapes large scale distribution patterns. In general, ocean productivity has a dome-shaped relationship with turbulence, a facet of the macro-scale that is echoed by the intimate experience of individual plankters.

The fifth chapter deals with motility, in particular how the motile behaviour of individuals is manifest at the population level. This is perhaps the most mathematical part of this work as it deals with the rarified topic of stochastic calculus – a rigourous means of analysis and modelling processes such as Brownian motion and random walks. As we will see, curious effects appear in population level distribution patterns that depend on the smallest details of how organisms enact their locomotion behaviour. In particular, how organisms react to external stimuli, changing their swimming speed or path geometry, lead to emergent population level phenomena that are either dispersive or aggregating.

This leads into a chapter on patchiness. That is, the distribution of solutes, tracers, particles and plankton in the marine pelagic environment is far from homogeneous. These distributions are often organized into patches and clusters at peak concentrations several times higher than mean background levels, and at scales ranging from millimetres to kilometres. Both physical (*e.g.* turbulent stirring) and biological (*e.g.* growth, grazing, migration, social-aggregation) processes are involved in the generation, maintenance and dissipation of patchiness.

The last 2 chapters deal with the implications of fitness-seeking behaviour, both at the individual level, and its consequences for population dynamics. A fairly robust (albeit incomplete) estimate of fitness following a given behaviour can be deduced from the organism's probability of surviving over a given time and its net energy acquisition over the same time interval. Encounter rates, sensing ability, motility, turbulence and patchiness all enter into these estimates. The aim here is to couple together elements of hydrodynamics, oceanography, evolutionary and behavioural ecology, rich in mechanistic detail, from which the interactions between plankton emerge – that is, a rationalization of their behaviour. Relatively simple arguments along these lines (*i.e.* the tradeoff between benefit, cost and risk) give insights as to why plankters behave the way they do: how fast they swim, their optimal path geometry, vertical positioning and time allocation in foraging. It also resolves some aspects of population dynamics – the paradoxes of omnivory and enrichment for instance – in that adaptive (*i.e.* fitness-seeking) behaviour has a stabilizing effect on community dynamics.

As a whole, this work promotes a new way of addressing marine ecosystem models, where the interactions between populations, rather than being prescribed from empirical observations, emerge from a mechanistic application of physical and evolutionary principles on individual members of these populations. The vision here is to utilize this extra information to lift marine ecosystem modelling out of its physico-chemical paradigm, to a more biologically oriented approach guided by the fundamental laws that govern all life.

This monograph has been prepared in partial fulfilment of my application for the degree of Doctor Technices at the Technical University of Denmark. As such, it focusses primarily on my own research. The document as a whole is drawn up so as to pull a red-tread through the last decade of my research activities. I have, however, also gone to some effort to provide a wealth of background material at a level that many be understandable a more general readership (*e.g.* graduate students, professionals) in marine science.

2 Encounter Rates

Nearly all important life processes of plankton, including growth, reproduction and mortality, are mediated by encounter rates – the rate at which individual plankters contact other organisms, particles or patches of resources. The rate at which a plankter can grow depends on its contact with food items, its reproduction rate depends in part on its contact with mates, while death is a likely outcome from an encounter with a predator. Encounter rates are the currency of planktonic interactions, and as such, play a fundamental role in ecology.

Encounter rates are determined by three basic parameter; the concentration (*i.e.* number per unit volume) of encounter partners (you can think of these as prey, but they can just as likely be mates, predators, particles or patches), the relative speed of the organism (searcher) and its encounter partner (target), and the distance at which an encounter takes place. The encounter distance depends on the sensory ability of the organisms involved (*e.g.* tactile, hydromechanical, chemical, or visual), and how the environment may attenuate these signals. This issue will be examined in more detail in a following chapter. The detection distance, and how it varies in space with respect to the positioning of both the organism detecting and that being detected, in general traces out a complex shape in space. The encounter volume for a drifting jelly fish for instance, may be approximated by a vertical cylinder, for a copepod nauplius a sphere, while that of a larval fish may be two cones projecting outwards form its eyes. Chemical clouds and plume present even more complex shapes. For the moment I will consider the simplest case; a spherical detection zone of fixed radius R, which provides a solid basis from which the effects of more complex encounter volume geometries may be assessed.

A theme in this chapter is the influence of path geometry on encounter rates – proceeding from linear paths (termed ballistic motion) through to more convoluted swimming tracks of a random walk nature. The aim is to present simple mechanistic descriptions of the relationships between behaviour, sensory ability and the rate at which individual plankters interact with each other and their resources.

2.1 Ballistic motion

2.1.1 Uniform ballistic motion of the predator

Perhaps the simplest formulation for encounter rates comes from considering a predator continually swimming in a straight line (termed ballistic motion) in a field of stationary prey. Further, assuming that the predator can detect prey at a given distance R irrespective of their relative direction, the problem becomes one of flux into a perfectly absorbing translating spherical collector. If C is the concentration of prey (*i.e.* number per unit volume), and v is the speed of the predator, then the number of prey entering through a small annulus on the surface of the sphere per unit time is the projection of the area of the annulus perpendicular

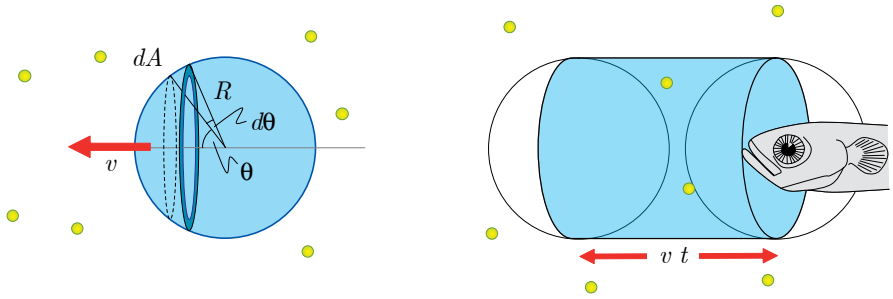


Figure 2.1: Integrating the flux of of particles into a moving sphere.

to the direction of motion times velocity times concentration. That is

$$dZ = CvdA\cos\theta = Cv2\pi R^2 \sin\theta\cos\theta d\theta \tag{2.1.1}$$

(*cf.* Fig. 2.1) where θ is the angle between the direction of motion and the annular position on the surface of the sphere. Integrating over all these angles for which the flux is into the sphere, *i.e.* from $\theta = 0$ to $\theta = \pi/2$ gives the total flux

$$Z = 2\pi R^2 Cv \int_0^{\pi/2} \sin\theta\cos\theta d\theta = \pi R^2 Cv \tag{2.1.2}$$

which is the encounter rate – number of prey contacted per unit time – that the predator experiences. It can easily seen from Eq. 2.1.2 that the total number of contacts in a time interval t is $Zt = C\pi R^2 vt$; this is the number of prey contained in a cylinder of radius R and length vt. Indeed the cartoon often used to illustrate encounter rates is that of a predator – a larval fish for instance – sweeping out a tube in space and collecting all prey within that tube.

2.1.2 Uniform ballistic motion of the prey

Of course, prey are oftentimes not stationary but move in their own right. Indeed, in some instances, the predator is motionless, laying in ambush, relying on the prey movement to bring about contacts. In the above case where a predator moves at a uniform velocity (*i.e.* fixed speed and direction) through a field of stationary prey, it is quite easy to see that the

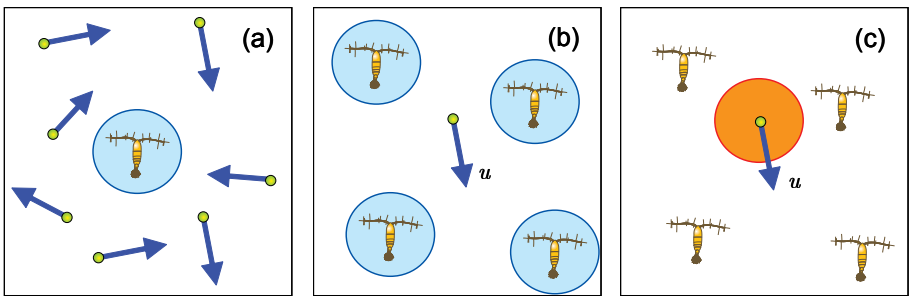


Figure 2.2: Randomly moving prey and a stationary predator; (a) as a predator would see it, (b) as a prey would see it, and (c) the equivalent moving "danger zone" view.

encounter rate is the prey concentration times the rate at which "new" fluid volume is swept out by the search volume.

What now, when the predator remains stationary, and the prey move, all with the same speed (call it u), but each with different random directions! At first this may seem extremely complicated, but it can be simplified by noting a certain symmetry in the encounter process. If we simply count the number of contact events per unit volume in a field of predators and prey, irrespective of how they move, then the rate at which contact events occur is $E = PC\beta$ where C is the prey concentration, P the predator concentration and β is the encounter kernel (dimensions: volume/time). The rate at which an individual predator encounters prey is

$$Z_P = E/P = C\beta$$

while the rate at which an individual prey is encountered by a predator is

$$Z_C = E/C = P\beta.$$

That is, the same encounter kernel enters into both expressions. Thus, instead of considering many randomly moving prey and a single predator, we can consider a single moving prey and many stationary predators (Fig. 2.2). Furthermore, because contacts occur at a distance R, we can consider the prey as travelling with a spherical "danger" zone – any time a predator enters this zone, a contact is made. The geometry of this is exactly the same as that considered in Sec. 2.1.1. Thus we can immediately write

$$Z_C = \pi R^2 P u \tag{2.1.3}$$

(the encounter rate of predators by a prey) from which it immediately follows that the encounter kernel is $\beta = \pi R^2 u$ so that

$$Z_P = \pi R^2 C u \tag{2.1.4}$$

This is the encounter rate experienced by an ambush predator in a field of ballistic prey, each of which is moving with uniform speed in uniformly random directions.

2.1.3 Non-uniform swimming speed.

Swimming speeds, in general, are not uniform. Some prey swim faster, some slower; and individual prey can change their swimming speed over time. Here I consider ballistic swimming where swimming speed is variable. Again, this is a factor that at first may appear cumbersome to incorporate in a simple model. However, it can be immediately noted that encounter rate is linear in prey concentration C and swimming speed u. For instance, we can consider the encounter rate with 2 populations of otherwise similar prey, a fast swimmer (speed u_f and concentration C_f) and a slow swimmer (speed u_s and concentration C_s).

The net encounter rate is simply the sum of the encounter rates with each sub-population:

$$Z = \pi R^2 \left(C_f u_f + C_s u_s\right) \tag{2.1.5}$$

Of course $(C_f u_f + C_s u_s)/(C_f + C_s)$ is simply the mean swimming speed of the whole population, and $C_f + C_s = C$ is the overall concentration of prey irrespective of their swimming speed. This can be generalised to any number of sub-populations with different swimming

speeds, or for that matter a continuous distribution $c(u)$. Thus, provided R remains constant, and with $C = \int c(u)du$,

$$Z = \pi R^2 C \langle u \rangle \tag{2.1.6}$$

where $\langle u \rangle$ is the mean swimming speed exhibited in the prey field given by

$$\langle u \rangle = \frac{1}{C} \int_0^\infty uc(u)du \tag{2.1.7}$$

2.1.4 Ballistic motion of predator and prey

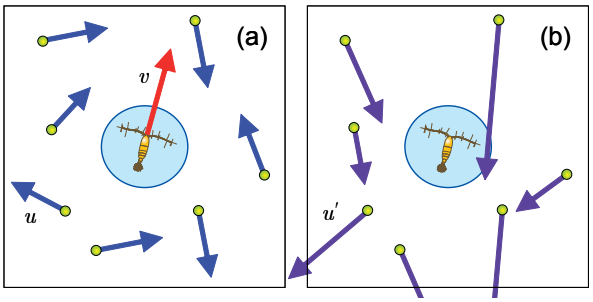


Figure 2.3: Moving predator and prey; (a) in the "laboratory" rest frame, and (b) in the predator rest frame.

I now approach the problem of the encounter rate when both predator and prey move. The first to address these situations were Gerritsen and Strickler (1977), who derived two slightly different formulations depending on whether the prey moves faster than the predator or vice-versa. Specifically

$$\begin{aligned} Z &= \pi R^2 C \left(\frac{u^2 + 3v^2}{3v} \right) \quad \text{for} \quad v > u \\ Z &= \pi R^2 C \left(\frac{v^2 + 3u^2}{3v} \right) \quad \text{for} \quad u > v \end{aligned} \tag{2.1.8}$$

The means by which this formulation is derived is not particularly transparent, and a much simpler derivation follows using Galilean invariance. For the case considered above where the prey is stationary, it should make no difference to the flux if the whole field, predator and prey move with some velocity $\mathbf{U}$ with respect to a fixed reference frame. The same applies for the case where both predator and prey move (Fig. 2.3). In particular, we can always find a reference frame in which the prey is stationary and the predator moves. The relative motion of predator and prey is invariant.

For a given situation where in the rest frame, the prey moves at an angle ϕ with respect to the direction of motion of the predator, the relative speed u' is given by the cosine rule (Fig. 2.4). That is

$$u' = \sqrt{v^2 + u^2 + 2uv\cos(\pi - \phi)} \tag{2.1.9}$$

where v and u are the speed of the predator and prey in the rest frame. The mean relative speed is then the average of u' over all possible relative angles:

$$\langle u' \rangle = \frac{1}{4\pi} \int_0^\pi \left(v^2 + u^2 + 2uv\cos(\pi - \phi) \right)^{1/2} 2\pi \sin\phi d\phi \tag{2.1.10}$$

which gives

$$\langle u' \rangle = \frac{|u+v|^3 - |u-v|^3}{6uv} \tag{2.1.11}$$

Since we can invoke translational invariance, *i.e.* move to a reference frame where the predator

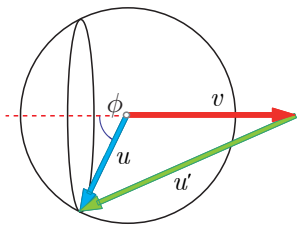


Figure 2.4: The relative velocity between a moving predator (speed v) and moving prey (speed u at a relative angle ϕ. The mean relative speed $\langle u' \rangle$ is got by integrating over all possible angles ϕ.

is at rest, the encounter rate for ballistically moving predator and prey, can the simply be got by inserting $\langle u' \rangle$ for $\langle u \rangle$ in Eq. (2.1.6). That is

$$Z = \pi R^2 C \frac{|u+v|^3 - |u-v|^3}{6uv} \tag{2.1.12}$$

This formulation is identical to that derived by Gerritsen and Strickler (1977) in that it does not presuppose anything of the relative dimension of u and v.

2.1.5 Gaussian ballistic motion – the Maxwell distribution

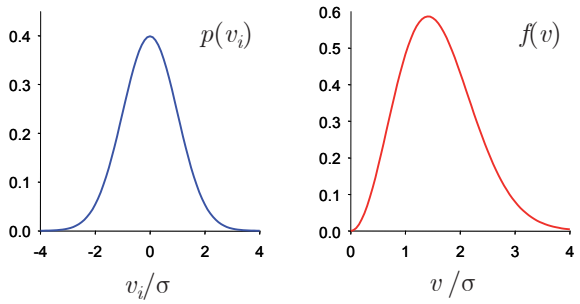


Figure 2.5: Gaussian distribution of a velocity component $p(v_i)$ and the corresponding Maxwell speed distribution $f(v)$.

Not all predators and prey move with the same speed. Neither do individuals always maintain the same speed. There is some variation in the swimming velocities of both predators and prey. I start by assuming that the 3 orthogonal components of the swimming velocity of the predator $\mathbf{v}$ have a Gaussian distribution with the same variance. That is, the swimming velocity is isotropic – displaying the same statistics in all directions. In particular, the probability of the organism swimming with an x-component of velocity (denoted v_x in the small interval $[v_x, v_x + dv_x]$ is given by:

$$p(v_x)dv_x = \frac{dv_x}{\sigma\sqrt{2\pi}} \exp(-v_x^2/2\sigma^2) \tag{2.1.13}$$

with similar expressions for the y- and z-components (*i.e.* v_y and v_z). Precisely the same form of expression can be written for the components of the prey swimming velocity $\mathbf{u}$, although in general with a different variance.

While Eq. 2.1.13 says something about the probability distribution of individual velocity components, the probability distribution for swimming speed, v has a somewhat different form:

$$f(v)dv = \frac{4\pi v^2 dv}{(\sigma\sqrt{2\pi})^3}\exp(-v^2/2\sigma^2) \tag{2.1.14}$$

where $v = \sqrt{\mathbf{v}\cdot\mathbf{v}}$. This is the Maxwell speed distribution[1], and its derivation from the Gaussian distribution of velocity components can be found in classic text books on kinetic theory. [2] One of the outcomes of these distributions is that the mean speed is given by

$$\langle v\rangle = \langle\sqrt{\mathbf{v}\cdot\mathbf{v}}\rangle = \sqrt{\frac{8}{\pi}}\sigma \tag{2.1.15}$$

while the root mean square velocity is

$$v_{rms} = \sqrt{\langle\mathbf{v}\cdot\mathbf{v}\rangle} = \sqrt{3}\sigma \tag{2.1.16}$$

Note the different order in which averaging and square root are applied in these formulae.

The point of interest here is that two Gaussian processes with different variance, when summed, produce a new Gaussian process with a total variance equal to the sum of the variance of the two original processes. In particular, if swimming speeds for the predator and prey have a Maxwell distribution (*i.e.* their velocities have a Gaussian distribution), then their relative speed will also have a Maxwell distribution. That is

$$\begin{aligned}(u'_{rms})^2 &= \langle\mathbf{u}'\cdot\mathbf{u}'\rangle = \langle(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})\rangle \\ &= \langle\mathbf{u}\cdot\mathbf{u}\rangle + \langle\mathbf{v}\cdot\mathbf{v}\rangle = (u_{rms})^2 + (v_{rms})^2\end{aligned} \tag{2.1.17}$$

Thus following from Eq. 2.1.6 it is almost trivial to write the result that

$$Z = \pi R^2 C\sqrt{\langle u\rangle^2 + \langle v\rangle^2} \tag{2.1.18}$$

This neat trick was first applied to encounter rates by Evans (1989). In my own work, I invariably use Eq. 2.1.18 rather than Eq. 2.1.12, both because it is simpler, and it address the variation of swimming speeds witnessed in nature. It should be noted that Eq. 2.1.18 is written in terms of mean swimming speed (*i.e.* as defined in Eq. 2.1.15). If root mean squared velocity (*i.e.* as defined in Eq. 2.1.16) is used instead, an additional factor $\sqrt{8/3\pi}$ will appear in Eq. 2.1.18. It can also be noted that while the formulae Eq. 2.1.12 and Eq. 2.1.18 differ in form, numerically they differ by at most 6%, the difference between 4/3 and $\sqrt{2}$. This is usually more than adequate given the uncertainties in other parameters involved (the perception distance R for instance).

2.2 Random walk encounters

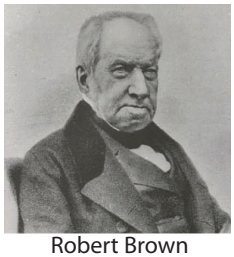
Up until this point I have considered only ballistic motion – where organisms continually swim along straight-lines. Some times this may be a reasonable approximation, at other times not. Many planktonic organisms from bacteria to copepods change their direction of swimming

[1] A general distribution of this type is more properly referred to as a Rayleigh distribution.

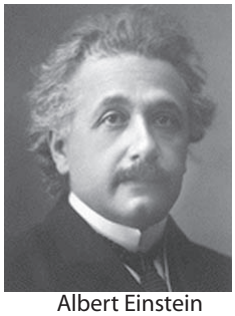
[2] The trick is to recognize that integrating over small volumes $d\mathbf{v}$ in velocity space is equivalent to integrating over thin spherical shells $4\pi v^2 dv$ where v is speed.

episodically. In many instances (but not always) these changes of direction appear to be random, with no preference with respect to either the spatial coordinates (*e.g.* up, down, left, right) or their previous direction of motion. This opens a whole new class of encounter processes, the kinematics of which are surprisingly, quite different from the ballistic encounter processes discussed so far.

2.2.1 Brown and Einstein


Robert Brown

If one looks at small inert particles in a fluid under a microscope, one will see these particles move and zigzag about in an apparent random fashion. This is so called Brownian motion. While Jan Ingenhousz (1730–1799) had observed the irregular motion of coal dust particles already in 1785, Brownian motion is generally regarded as having been discovered by the Scottish botanist Robert Brown (1773–1858) in 1827. Hence the name. While closely examining pollen grains suspended in water, he noted that they were "filled with minute particles" that were "very evidently in motion" zipping about in jittery movements. To his credit, Brown forwent the obvious explanation, that some vital force was involved. He remained curious and repeated the experiment to observe the same jittery motion for dust particles, (crushed rock from the nose of the Sphinx for instance) - material that was clearly dead. Brownian motion was evidently, a physical phenomenon.


Albert Einstein

Some of the mathematics describing the kinematics of Brownian motion was formulated already in 1880 by the Danish astronomer Thorvald N. Thiele (1838–1910) and the French mathematician Louis Bachelier (1870–1946) in 1900 in his PhD thesis "The theory of speculation". A physical explanation however, was first provided by Albert Einstein (1879–1955) "Investigations on the theory of Brownian movement" in *Annalen der Physik* published in his extraordinary year 1905.

In Brownian motion, Einstein saw the fingerprint of atomic and molecular motion. In the late 19^{th} century, thermodynamics was a hot topic (pardon the pun) – the crowning achievement of which was the kinetic theory of matter expressed in the work of James Clark Maxwell (1831–1879) and Ludwig Boltzmann (1844–1906). If the kinetic theory was correct *i.e.* that heat is a measure of the kinetic energy of the random and continual motion of molecules and atoms, then small particles embedded in a fluid will be subject to continual and random bombardments, causing these small particles to move also. Up until the early 20^{th} century, kinetic theory was just a theory — it was consistent at the macro-scale, but still lacked empirical evidence at the molecular level. It was Einstein and those who followed him: Paul Langevin (1872–1946), Marian Smoluchowski (1872–1917) and Jean Baptiste Perrin (1870–1942), who provided this evidence. It was Einstein for instance, who came up with a kinetic theory of diffusion, namely that the diffusivity of a molecule is given by:

$$D = \frac{2RT}{Af} \tag{2.2.1}$$

where R is the universal gas constant, T the absolute temperature, f the coefficient of friction of the fluid, and A Avogadro's constant. It was Perrin who was able to estimate Avogadros, that truly number that relates the number of molecules (or atoms) in a mole (6×10^{23}), demonstrating the physical reality of atoms.

I include this little excursion as more than just a historical note. There is an important concept here that can be carried over into plankton ecology. That is, that macro-scale phenomena (in ecology this is the ecosystem, population, or community while in physics it is the continuum – a gas or a liquid for instance) can be both deduced from and give insight into processes that are played out at the scale of individuals (organisms in ecology, molecules in physics). In the same way that Brownian motion provides a kinetic theory of diffusion, random walks can be used to describe the dispersal and migration of populations based on the kinesis and taxis of their individual members.

2.2.2 Diffusive nature of random walks

Uniform jumps and time steps: The most accessible avenue to understanding the workings of random walks is to consider a discrete jumper in one dimension; it can jump either to the left or the right in a distance λ per time step. That is, its position x from one time step to the next is

$$x_{i+1} = x_i \pm \lambda \tag{2.2.2}$$

Now if we consider a large number of these random jumpers, all obeying the same random process, not interfering with each other, then the mean and variance of the distribution after 1 time step is

$$\langle x_{i+1} \rangle = \langle x_i \rangle \tag{2.2.3}$$

$$\langle (x_{i+1})^2 \rangle = \langle (x_i)^2 \rangle + \lambda^2 \tag{2.2.4}$$

That is, at each time step, the mean displacement remains the same but the variance increases by λ^2. Projecting backwards in time, if all the particles were at $x = 0$ at $t = 0$, after n time steps of duration δ, *i.e.* at time $t = n\delta$, the mean position of the particles is still

$$\langle x(t) \rangle = 0 \tag{2.2.5}$$

while the variance

$$\langle x(t)^2 \rangle = n\lambda^2 = \frac{\lambda^2}{\delta} t \tag{2.2.6}$$

The variance says something about the shape of the distribution, and it grows linearly in time. The time rate of change of variance is the diffusivity, D. That is, for the uniform jump-length, uniform time-step example above, this becomes

$$D = \frac{1}{2} \frac{d}{dt} \langle x(t)^2 \rangle = \frac{\lambda^2}{2\delta} \tag{2.2.7}$$

The precise diffusive nature of a random walk can be derived by observing the number of possible paths by which a random walker can get to a particular location in a given time. If, as above, the time steps and jump lengths are discrete, this is given by the binomial distribution, which in the limit of many time steps, can be approximated by a Gaussian distribution. There are any number of classic texts where this can be found (*e.g.* Berg, 1992; Okubo and Levin, 2001) and will not be repeated here – suffice to say that in this limit, the probability of finding a random walker in a small space interval ($[x, x + dx]$) and time interval ($[t, t + dt]$) is given by

$$p(x, t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(\frac{-x^2}{4Dt}\right) \tag{2.2.8}$$

Variable run lengths and time steps: This can be very generally extended for the case where jump-lengths and time-steps are variable (but still in 1 dimension). In this case

$$D = \frac{\langle \lambda^2 \rangle}{2\langle \delta \rangle} \tag{2.2.9}$$

This can be seen from Eq. 2.2.6 where instead of the ensemble average, we first average over all n time steps, thus

$$\overline{x(t)^2} = n\overline{\lambda^2} \tag{2.2.10}$$

where the overbar indicates averaging over time, compared to $\langle \cdots \rangle$ which indicates averaging over many realizations of the same process. While the number of jumps for each estimate of variance is equal (n), the time over which each realization is estimated may vary: $t' = n\bar{\delta}$. The mean time for all realizations is thus $t = \langle t' \rangle = n\langle \bar{\delta} \rangle$, but since the process is assumed stationary, the statistics of δ should be the same irrespective of whether an average is taken over many time steps for the same realization, or if the average is taken for the same time step but for many different realizations. That is $\langle \bar{\delta} \rangle = \langle \delta \rangle = \bar{\delta}$. This is known as an ergodic process. Thus Eq. 2.2.9 follows directly form Eq. 2.2.10.

I now change focus to random walks that have a somewhat more relevance for planktonic organisms – swimming rather than jumping – that is, I assume the organism swims at a constant speed u, and changes direction at time intervals that may vary. The "jump" length thus relates to the swimming speed and the time interval: $\lambda = u\delta$. A prototypical example of this type of motility is that displayed by bacteria, from which we can borrow the terminology: λ is the run-length, and δ is the tumble period (= tumble rate)$^{-1}$. Thus Eq. 2.2.9 can be written

$$D = u^2 \frac{\langle \delta^2 \rangle}{2\langle \delta \rangle} \tag{2.2.11}$$

I now examine the case where the tumble period has a specific distribution: the exponential distribution. That is, the probability of the organism changing its direction in a small time interval Δt is $1 - \exp(-\Delta t/\tau)$ where τ is the intensity of tumble events. That is, the probability of tumbling in a short interval is the same irrespective of anything that has happened before.

Of particular interest is the probability distribution of δ, the time interval between tumble-events following this rule. The probability of a tumble occurring in a small time interval $\delta + \Delta\delta$, is equal to the probability of it *not* occurring up to time δ, and then occurring in the subsequent interval $\Delta\delta$. Remember that the probability of something not occurring = 1 minus the probability that it *does* occur. That is, the portability of *not* changing direction up to time δ is $\exp(-\Delta\delta/\tau)^n = \exp(\delta/\tau)$ where $\delta = n\Delta\delta$. Thus

$$p(\delta) = \lim_{\Delta\delta \to 0} \exp(-\delta/\tau)(1 - \exp(-\Delta\delta/\tau) = \exp(-\delta/\tau)\frac{d\delta}{\tau} \tag{2.2.12}$$

Therefore, it follows that

$$\langle \delta \rangle = \frac{1}{\tau}\int_0^\infty \delta e^{-\delta/\tau} d\delta = \tau \tag{2.2.13}$$

and

$$\langle \delta^2 \rangle = \frac{1}{\tau}\int_0^\infty \delta^2 e^{-\delta/\tau} d\delta = 2\tau^2 \tag{2.2.14}$$

That is τ, the intensity of the exponential process is simply the mean tumble period, and variance of the tumble periods is twice the square of the mean tumble period. This factor of 2 will appear again in relating the mean variance to the mean intensity of an exponential process. Thus, in one dimension, the diffusivity of an exponentially distributed run-tumble process is $D = u^2\tau$.

A 3 dimensional view: It is almost trivial to extend these arguments to 3 dimensions. We can replace $x(t)$ and λ with the 3-vectors, $\mathbf{x}(t)$ and $\boldsymbol{\lambda}$, and proceed along the same lines to give

$$\begin{aligned}\langle \mathbf{x}(t) \cdot \mathbf{x}(t) \rangle &= \langle \boldsymbol{\lambda} \cdot \boldsymbol{\lambda} \rangle \frac{t}{\delta} \\ &= \left(\langle \lambda_x^2 \rangle + \langle \lambda_x^2 \rangle + \langle \lambda_x^2 \rangle \right) \frac{t}{\delta} = 3\langle \lambda^2 \rangle \frac{t}{\delta}\end{aligned} \tag{2.2.15}$$

where in the last step we have assumed spatial invariance; that the statistical characteristics of runs and tumbles are independent of direction. Given this, a consistent definition of diffusivity then becomes

$$D = \frac{1}{2N} \frac{d}{dt} \left\langle |x(t)|^2 \right\rangle \tag{2.2.16}$$

where N is the number of dimensions (1, 2 or 3 in the normal world).

I said almost trivial above. In one dimension, the meaning of a random direction is quite clear. At a tumble event, the organism can either reverse or continue on in the same direction, each with 50-50 probability. At higher dimensions however, there is a continuum of directions to choose from. Technically, in order for Eq. 2.2.15 to be true, run directions over all relative lags must be non-correlated. That is

$$\langle \boldsymbol{\lambda}_i \cdot \boldsymbol{\lambda}_{i+j} \rangle = 0 \quad \forall j \neq 0 \tag{2.2.17}$$

Generally it suffices to require that $\langle \cos\theta_i \rangle = 0$ where θ_i is the angular change between successive runs. Leaving the details until later, for the moment we note that when $\langle \cos\theta \rangle = c$, *i.e.* there is a correlation between successive runs, diffusivity may be written as

$$D = \frac{1}{N} \frac{u^2\delta}{1-c}$$

2.2.3 Diffusive encounter rates

An encounter model can thus be constructed proceeding from the assumption that the prey executes a random walk and is therefore diffusive. If we consider an ambush predator (*e.g.* the copepod *Acartia tonsa*) then the flux of random walking prey (*e.g.* the ciliate *Balanion comatum*) to it can be treated exactly analogously to the flux of diffusing material to a perfectly absorbing sphere.

If $\rho(r)$ is the concentration of prey as a function of distance r from the centre of the sphere, then

$$\frac{\partial \rho}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(r^2 D \frac{\partial \rho}{\partial r} \right) \tag{2.2.18}$$

subject to the boundary conditions that $\rho(R) = 0$ (that is, all prey is consumed at the surface of the predator's capture zone), and $\rho(r \to \infty) = C$, that at distances far from the predator,

the prey concentration approaches a constant background value. The prey flux (*i.e.* the encounter rate) is then given by

$$Z(t) = -\oint D\frac{\partial \rho}{\partial r}\hat{\mathbf{r}} \cdot d\hat{\mathbf{s}} \tag{2.2.19}$$

which gives, using $D = u\lambda/3$, the time dependent solution

$$Z(t) = \frac{4}{3}\pi CRu\lambda\left(1 + \sqrt{\frac{3R^2}{\pi u\lambda t}}\right) \tag{2.2.20}$$

While this solution may look pathological – infinite flux as $t \to 0$ – the total number of prey encountered is well behaved, varying as $t^{1/2}$ for small t. As t increases, the flux (*i.e.* encounter rate) tends to

$$Z(t) = \frac{4}{3}\pi CRu\lambda \tag{2.2.21}$$

Note that $\lambda = u\tau$, so that Eq. 2.2.20 and 2.2.21 can also be written in a slightly different forms involving τ rather than λ.

2.3 Ballistic – diffusive transition

We thus have two different encounter rate models, both based on apparently sound physical reasoning, that give qualitatively and quantitatively different results. Consider for instance an ambush predator, ($v = 0$), the ballistic models, Eq. 2.1.2, Eq. 2.1.4 and Eq. 2.1.6 all give

$$Z = \pi R^2 Cu. \tag{2.3.1}$$

while if we recognize the random motility of the prey motility from the outset, we would predict

$$Z = \frac{4}{3}\pi CRu^2\tau \tag{2.3.2}$$

Not only do these give different numerical estimates, they have a different dependence on governing parameters: R^2 for the former but R for the latter, u^2 for the latter but u for the former. Clearly, both these formulae cannot be correct at the same time.

2.3.1 Ballistic and diffusive encounter rates – a paradox?

To get some idea of how these estimates fit together, I examine the process using a numerical simulation. Such simulations are actually quite simple to implement and can illustrate processes much more readily than rigourous mathematical analyses. Consider for instance a spherical capture zone of $R = 1$ mm radius centred in a 1 L volume cube (Fig. 2.6). I seed the volume with a number of particle that move with a uniform speed, $u = 1$ mm/s, and tumble to a new direction (uniformly random in 3 dimensions) at a mean time interval τ. That is, the probability of tumbling to a new direction in a time step Δt is $1 - \exp(-\Delta t/\tau)$. The encounter rate can then be estimated from the number of particles that enter the capture zone each time step.

Results from 2 model runs are presented in Fig. 2.7. In these, the swimming speed and perception distance are the same, but the tumble intervals, τ , are 0.1 s and 2 s respectively.

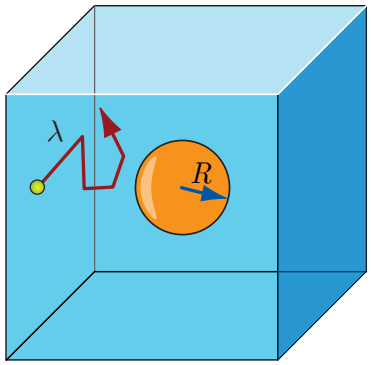


Figure 2.6: Simulation geometry to examine the ballistic – diffusion transition. How does the particle flux into the sphere depend on the relative size of the sphere, R and the mean run-length λ? Particle move according to a run-tumble motion; those entering the capture zone continue to move but are set to "inactive" and don't count towards any future encounters. Particles passing out of the simulated region are reintroduced through the opposite side, and any of these that were inactive are reactivated.

The estimated encounter kernel is given by $\beta = N_{enc}V/(N_{act}\Delta t)$ where N_{enc} is the number of encounters in the time interval Δt, and N_{act} is the number of active particles in the simulation volume V.

When $\tau = 2$ s (Fig. 2.7.b), the measured encounter kernel conforms well to the ballistic estimate, and remains constant over time. This may be expected since $\lambda > R$. Using the diffusive estimate in this case would over estimate the realized encounter rate by a factor 2.5. For $\tau = 0.1$ s (Fig. 2.7.a), the measured clearance rate is initially ballistic, but falls off towards the diffusive limit over time. The fall off rate appears to be well described by the time dependent diffusive estimate. In this case, the over all clearance rate is diffusive in nature as may be expected since $\lambda < R$. Using the ballistic estimate in this case would over estimate the realized encounter rate by a factor of up to 7.

Ballistic encounters are inherently more efficient than diffusive encounters. The former applies at small and the latter at larger spatio-temporal scales; specifically when the motility length scale $\lambda = v\tau > R$, the encounter rate is determined by the ballistic formulation at all times, whereas when the motility length scale $\lambda = v\tau < R$, the encounter rate is initially determined by the ballistic formulation, but tends towards the diffusive formulation over a time scale R^2/D.

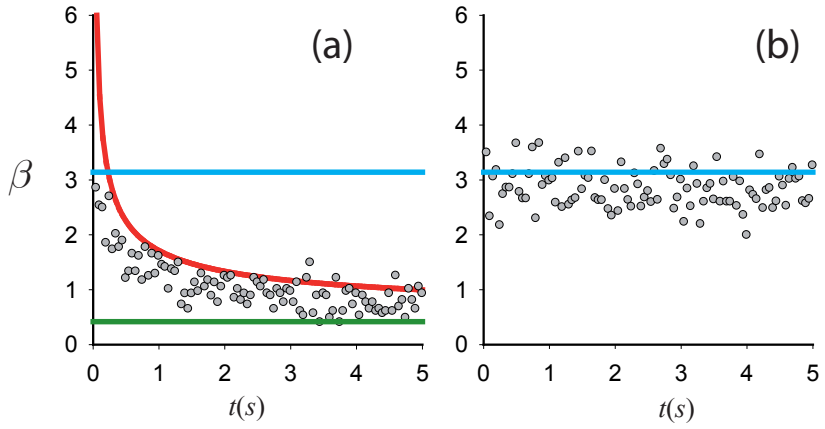
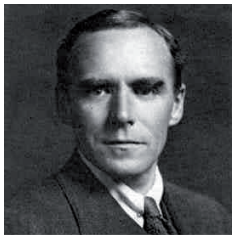


Figure 2.7: The clearance rate β as a function of time for (a) $\lambda = 0.1R$ and (b) $\lambda = 2R$. Data from numerical simulations and the ballistic and diffusive estimates are shown.

2.3.2 Continuous random walk

To bring the ballistic – diffusive transition into a more rigourous frame work, consider now a continuous random walk. Many planktonic organisms for instance, move in a way that is continuous and sinuous, either by their own swimming or by turbulent motions of the water. While not being composed of discrete straight-line displacements interspersed with random changes of direction, these motility patterns can still have all the hall marks of a random walk.

Consider for instance an organism that swims at a uniform speed u, along a continuous curvy trajectory obeying some fixed set of rules. What is the expectation value of the net displacement of the organisms over time? Specifically, how does the expectation value of the net distance travelled, ℓ, vary in time, or equivalently, with respect to the gross distance travelled L?


G.I. Taylor

The following analysis addressing this question is essential that followed by GI Taylor [3] in his seminal 1921 work in determining diffusion caused by continuous motion such as turbulence. Here it is presented in a more general 3 dimensional setting, and for a random walk motility.

The trajectory of a specific particle can be thought of a finite number (n) short sections $\boldsymbol{\xi}_i$, each of expectation length $\langle \boldsymbol{\xi}_i \rangle = \Delta$. The vector sum of all these short sections adds up to the net straight line displacement of the particle over the n time steps. The square of the net distance travelled is thus

$$\begin{aligned}
\ell^2 =& \langle \mathbf{x} \cdot \mathbf{x} \rangle \\
=& \langle (\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + \cdots + \boldsymbol{\xi}_n) \cdot (\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + \cdots + \boldsymbol{\xi}_n) \rangle \\
=& \langle (\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 \cdot \boldsymbol{\xi}_2 + \cdots + \boldsymbol{\xi}_n \cdot \boldsymbol{\xi}_n + \cdots \\
& \cdots + 2\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_2 + 2\boldsymbol{\xi}_2 \cdot \boldsymbol{\xi}_3 + \cdots + 2\boldsymbol{\xi}_{n-1} \cdot \boldsymbol{\xi}_n + \cdots + 2\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_n) \rangle
\end{aligned} \tag{2.3.3}$$

The first term is the mean square of the length of the trajectory:

$$\langle \boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 \cdot \boldsymbol{\xi}_2 + \cdots + \boldsymbol{\xi}_n \cdot \boldsymbol{\xi}_n \rangle = n\Delta^2 \tag{2.3.4}$$

The additional terms in Eq. (2.3.3) (there are $2(n-m)$ of them) are of the form $\boldsymbol{\xi}_i \cdot \boldsymbol{\xi}_{i+m}$ where m is the number of time steps between the sections. Each of these sections will have an equivalent expected length, so that the correlation between them depend on the cosine of the total angle between these vectors. That is

$$\langle \boldsymbol{\xi}_i \cdot \boldsymbol{\xi}_{i+m} \rangle = \Delta^2 c^m \tag{2.3.5}$$

Here I have assumed two things. Firstly that the expectation value of $\langle \cos\theta_i \rangle$ is a fixed value c, and secondly, that there is no correlation between the direction of the displacement steps at time off-sets greater than 1. Essentially, this means that the angles are independently drawn from a Gaussian distribution with mean 0 and variance $-2\ln(c)$. Combining these gives

$$\ell^2 = n\Delta^2 + 2\Delta^2 \left((n-1)c + (n-2)c^2 + \cdots + c^{n-1} \right) \tag{2.3.6}$$

[3] Sir Geoffrey Ingram Taylor (1886 – 1975) was a British physicist whose work features large in this monograph, both in his contributions to turbulence theory and in the field of the biophysics of propulsion by micro-organisms. His name, for instance is associated with the Taylor micro-scale, a fundamental measure of the fine structure of turbulence, the Taylor column, the trapping of geophysical fluid flows over bathymetric features, and the Taylor number, a dimensionless quantity relating the rotational inertia of a fluid to viscous forces. He is credited with being one of the most influential contributors to the field of fluid mechanics in the 20th century.

which contains an arithmetic-geometric series that can be simplified to give

$$\ell^2 = u^2\delta^2\left(n + \frac{2nc}{1-c} - \frac{2c(1-c^n)}{(1-c)^2}\right) \tag{2.3.7}$$

The number of segments is simply the total time divided by time taken to swim along each section. That is $n = t/\delta$ and, $\Delta = u\delta$. This gives

$$\frac{\delta}{1-c} = \tau \tag{2.3.8}$$

I now let $\delta \to 0$, noting the above and its consequences, namely that $\lim_{\delta\to 0} c = 1$, and $\lim_{\delta\to 0}(1 - c^{t/\delta}) = 1 - e^{-t/\tau}$, and get

$$\ell^2 = 2u^2\tau\left(t - \tau(1 - e^{-t/\tau})\right) \tag{2.3.9}$$

This is the central result of Taylors analysis. It shows that at short time scales (specifically for $t \ll \tau$), the net displacement grows linearly in time as

$$\ell \approx ut \tag{2.3.10}$$

which is precisely what we would expect if the particle (organism) were travelling (swimming) ballistically. At the other extreme, for $t \gg \tau$,

$$\ell^2 \approx 2u^2\tau t \tag{2.3.11}$$

That is the expectation value of the net distance squared (*i.e.* variance) grows linearly in time, a characteristic of a diffusive process.

2.3.3 Characterizing plankton motility patterns

Planktonic organisms exhibit a wide variety of swimming patterns from relatively erratic behaviours such as run - tumble and hop - sink, to more continuous swimming along nearly straight, helical, or sinuous paths. Examples of plankton motility patterns described in the literature include those of bacteria (Berg and Brown, 1977; Mitchell et al., 1995; Kiørboe et al., 2002), protists (Kamykowski et al., 1992; Fenchel and Blackburn, 1999; Bartumeus et al., 2003; Buskey and Stoecker, 1988; Kiørboe et al., 2004), copepod nauplii (Buskey et al., 1993; Titelman and Kiørboe, 2003b), and copepods (Schmitt and Seuront, 2001; Doall et al., 2002). Irrespective of the details of these swimming patterns, all such paths have a common characteristic in that at very small scales, the path tends to remain in a constant direction whereas at larger scales, a degree of randomness appears in the path direction. Paths are characterized by short term coherency but long term stochasticity. Specifically, this means that if we take ℓ as the straight-line displacement of an animal then for small time scales, the path is linear and

$$\ell \approx ut \tag{2.3.12}$$

where u is the mean speed of swimming. That is, the straight-line displacement increases linearly with time, a characteristic known as ballistic (Fig. 2.8). On the other hand, at longer times scales as the path becomes more convoluted, the straight-line displacement (net distance travelled) increases at a slower rate as more time is spent "double-backing" over areas already covered. At these large scales, we can write

$$\ell \propto t^{1/\alpha} \tag{2.3.13}$$

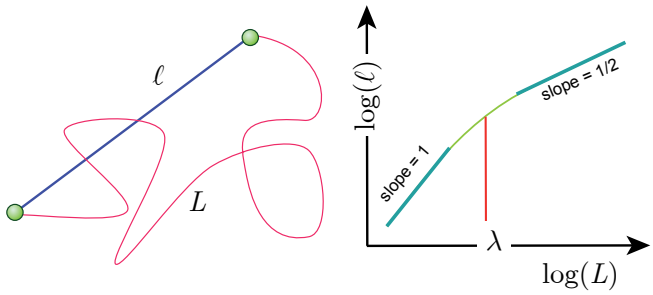


Figure 2.8: The gross displacement L and the net displacement ℓ and the gross-to-net displacement relationship for a continuous random walk on a log-log plot. The transition form ballistic to diffusive occurs at the correlation length scale λ.

where α is the fractal dimension, a coefficient in this case greater than 1. When the kinematics of the path trajectory conform to a Brownian random walk, then $\alpha = 2$, and the large scale dependence of ℓ is diffusive, $\ell \propto \sqrt{t}$. These relationships can equally well be expressed in terms of the gross distance travelled, L. For ballistic motion the net and gross distances travelled are identical, *i.e.* , $\ell = L$, whereas for diffusive motion $\ell \propto \sqrt{L}$. A log-log plot of ℓ versus L or t would reveal two distinct regions; one where the slope is 1, and another where the slope becomes 1/2. The temporal scale at which this transition happens is the "correlation time scale" (τ), and the corresponding spatial scale is the "correlation length scale" ($\lambda = u\tau$). These scales are of more than academic interest since, as we shall see later; they have implications for the rates at which motile organisms encounter their prey as well as their predators. These are precisely the characteristics displayed in Taylor's analysis of continuous random walks. We can in fact rewrite Taylor's formula (Eq. 2.3.9) in terms of gross and net distance travelled as:

$$\ell^2 = 2\lambda \left(L - \lambda(1 - e^{-L/\lambda}) \right) \tag{2.3.14}$$

The displacement versus time relationships for a wide variety of organisms form bacteria to adult copepods conform reasonably well to Taylors model, Eq. 2.3.9.

This behaviour has profound influence on the benefit-risk tradeoff for planktonic organisms (Visser and Kiørboe, 2006; Visser, 2008), an issue that will be explored in greater detail in Sec. 7.3.2.

2.3.4 The telegraph equation

For most practical applications in natural sciences, the diffusive approximation for Brownian motion is totally adequate. This is because the spatial scale of the process under consideration (*e.g.* the size of the source or sink region) is usually many times greater than the typical length of the Brownian runs. One can think of molecular diffusion: the mean free path of a molecule in water at room temperature is about 10^{-10} m, whereas the process scale (*e.g.* dissolved organic carbon leaking from a phytoplankton cell) is greater than 1μ m. However, for planktonic interactions, these scales are not well separated, and the transition from ballistic to diffusive processes may be important. The transition zone is sometimes termed meso-diffusive, and the governing dynamics can be written to first order as the telegraph equation (Goldstein, 1963; Okubo, 1980; Turchin, 1998; Uchaikin and Saenko, 2001).

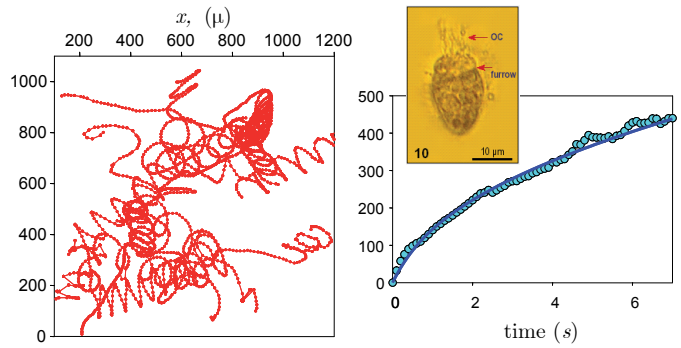


Figure 2.9: Analysing the track of a the ciliate *Balanion comatum* (inset (b)). (a) a 2-D projection of many trajectories, and (c) the net displacement vs time relationship; observations and the best-fit of Taylor's continuous random walk model. Motility parameters from this analysis yield $v = 220 \pm 10\mu/s$, $\lambda = 65 \pm 10\mu$, $\tau = 0.3 \pm 0.03s$, and $D = (50 \pm 10) \times 10^{-6} cm^2/s$.

To see there is something odd about diffusion at these scales, consider a population of 1-D random walkers (characteristics; speed v, correlation length scale λ) initially concentrated at a particular location $x = 0$. Over time, they will diffuse at a rate $D = v\lambda$, and, as we have seen before, their probability density function in the limit of many run-tumble cycles is given by:

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(\frac{-x^2}{4Dt}\right) \tag{2.3.15}$$

However, there is something not quite right with this prediction. It says that the probability of a random walker being at some large distance from $x = 0$ is small but finite. This is in conflict with what we know to physically possible; one of our random walkers cannot possibly be a distance from the origin greater than vt at a give time t. That is we require that $p(|x| > vt, t) \equiv 0$.

A formulation that tries to grapple within this is the telegraph equation. In its original incarnation, the telegraph equation related to a description of the the voltage and current on an electrical transmission line with distance and time, a topic of immediate practical value in the 1850's with the laying of the first Trans-Atlantic telegraph cable. It was first derived by William Thomson[4], and developed further by Oliver Heaviside [5].

In terms of variables already encountered, the telegraph equation can be written:

$$\frac{\partial^2 p}{\partial t^2} + \frac{2}{\tau}\alpha\frac{\partial p}{\partial t} = v^2\frac{\partial^2 p}{\partial x^2} \tag{2.3.16}$$

It describes composite dynamics that tend to the wave equation at short time scales (like ripples spreading from a stone tossed into a pond), and the diffusion equation at longer time

[4]William Thomson (1824 – 1907), a British physicist, mathematician and engineer. In 1892 he gained a peerage, and is known to history as Lord Kelvin. He is perhaps best known for his work on thermodynamics (absolute zero and the Kelvin temperature scale) although his interests ranged far and wide; from geology to marine science to fluid mechanics and beyond.

[5]Oliver Heaviside (1850 – 1925), British physicist, mathematician and inventor. A one-time telegraph operator in Denmark, his name features in "The Journey to the Heaviside Layer" in Andrew Lloyd Webber's musical Cats, an allusion to a sub-strata of the earth's ionosphere.

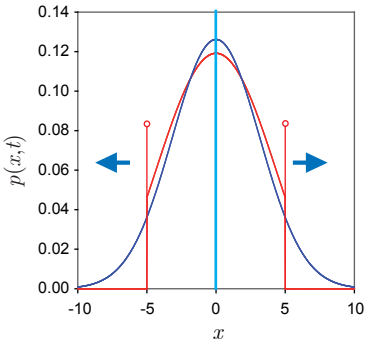


Figure 2.10: Solution to the telegraph equation (in red) compared to the diffusive solution (in blue) for a random walking particle initially at $x = 0$. The telegraph solution spreads out as two pulses travelling at speed v that gradually decay at a rate $e^{-\lambda t/2}$. At any given time, the probability density function is between $x = \pm vt$ is given by Bessel functions, which in the limit as $t \gg 1/\lambda$ approaches a Gaussian distribution.

scales. The form can be readily derived for a random walk (we will do this in one of the following chapters - once some technical issues have been addressed). The solution for an initial probability density function $\delta(x, 0)$ is given by

$$p(x,t) = p_{\text{int}}(x,t) + \frac{1}{2}e^{-t/\tau}\left(\delta(x-vt) + \delta(x+vt)\right) \tag{2.3.17}$$

where δ is the Dirac delta function. The interior solution $p_{\text{int}}(x,t)$ is given by

$$p_{\text{int}}(x,t) = \begin{cases} \dfrac{e^{-t/\tau}}{v\tau}\left(I_0(\xi) + \dfrac{t}{2\tau\xi}I_1(\xi)\right) & \text{for} \quad |\xi| < vt \\ 0 & \text{for} \quad |\xi| \geq vt \end{cases} \tag{2.3.18}$$

where I_0 and I_1 are modified Bessel functions of the first kind, and where

$$\xi = \frac{1}{2v\tau}\sqrt{v^2t^2 - x^2} \tag{2.3.19}$$

This is precisely the form we would expect from our physical intuition – pulses travelling outwards at a speed v with a diffusive like interior, Fig. 2.3.17. Having $p(x,t)$, it is thus possible to calculate the moments of the evolving distribution by integrating

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x,t)dx \tag{2.3.20}$$

While not straight forward, this can be done, and the result for the second moment (*i.e.* variance) becomes:

$$\langle x^2 \rangle = 2\left(v\tau\right)^2\left(e^{-t/\tau} - 1 + t/\tau\right) \tag{2.3.21}$$

which we can recognize as Taylor's formula Eq. 2.3.9. There is no 3-D analogue of this equation, and an analysis of the ballistic-diffusive transition along this tack remains elusive.

2.3.5 The Wiener sausage

The volume swept out by a spherical shell that is moving through space following a random walk trajectory, conforms to a mathematical entity known as the "Wiener sausage" (Fig. 2.11), named after the celebrated mathematician Norbert Wiener[6]. This turns out to be a

[6]Norbert Wiener (1894 – 1964), and American mathematician, who contributed to the theory of stochastic processes, electrical engineering and control theory. Principally a mathematician, he seems to have derived much of his inspiration from physical problems such as Brownian motion, that led to his seminal contribution to stochastic calculus in what is now known as a Wiener process. I will touch on this subject later in this monograph.

useful measure to number of real world processes in physics, chemistry and communication (*e.g.* Oshanin et al., 1994; Yang et al., 2000; Kesidis et al., 2003), as well as for the searching strategies of planktonic organisms (Viswanathan et al., 1999; Levandowsky et al., 1988).

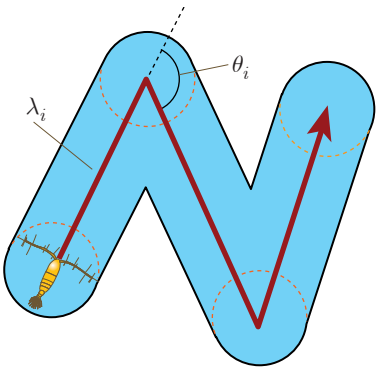


Figure 2.11: The Wiener sausage as traced out by a swimming copepod. A question of some relevance for encounter processes is how fast does the volume of the sausage increase as a function of the correlation length scale of the path, and the radius of the search volume.

While I have illustrated this in Fig. 2.11 as a sequence of straight-line segments interrupted by discrete turn events a zigzag path the treatment is formally identical for a continuous random walk (Taylor 1921) a meandering path. In either case, the random walk can be characterized by the correlation length scale λ If we consider an organism executing a random walk motility with search radius R, then its search volume equates to a Wiener sausage. It is clear that the relative size of λ and R has an immediate impact on the search efficiency of the organism.

At large times ($t \gg \tau$) and in the diffusive limit ($R \gg \lambda$), the expectation value of the volume (of a Wiener sausage (*i.e.* new volume scanned) is linear in time, and can be written as $V(t) = 4\pi\lambda Rvt/3$ (Berezhkovskii et al. 1989; Kesidis et al. 2003). This can be compared to the gross volume swept out $V(t) = \pi R^2 vt$. That is, the ratio of net to gross volume swept out goes as λ/R, so that the efficiency is relatively small for large R. This analysis is not particularly revealing however, as it relies on the diffusive assumption; that the search radius is much larger than the motility length scale. On the contrary, it appears that the motility length scale exhibited by a planktonic organism is smaller than its detection distance for prey while being of the same order as the distance at which a predator detects it (Visser and Kiørboe, 2006).

A full and rigourous examination of the characteristics of the Wiener sausage over the relevant parameter space, (*i.e.* as the motility length-scale varies with respect to the detection distance from $\lambda < R$ to $\lambda >> R$) has, as far as I know, never been achieved. As an indication, however, we can examine the asymptotic limits. For $\lambda \ll R$, the diffusive limit can be assumed to apply, giving

$$\beta_{\text{diff}} = \frac{4}{3}\pi R\lambda v \tag{2.3.22}$$

where β is the rate of change of the Wiener sausage *i.e.* maximum clearance rate. At the other extreme where $\lambda \gg R$, the ballistic limit applies

$$\beta_{\text{ball}} = \pi R^2 v \tag{2.3.23}$$

A simple relationship that bridges these equations for all ranges of λ and R is

$$\beta(\lambda, R) = \frac{dV}{dt} = \frac{4\lambda}{4\lambda + 3R}\pi R^2 v \tag{2.3.24}$$

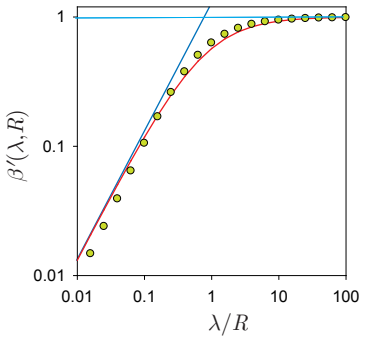


Figure 2.12: The normalized flux of Brownian particles to a spherical absorber as a function of correlation length scale and radius of the absorber: symbols indicate the approximation provided by Harris (1982). The red line is the analytic approximation Eq. (2.3.24) with the asymptotic values for the diffusive and ballistic limits indicated in blue. The invariance of encounter process suggests that this is the same as rate of increase of the Wiener sausage.

While the argument supporting this is somewhat circumstantial, a similar (albeit more complex) expression was more rigorously derived by Harris (1982) for the flux of Brownian particles into a perfectly absorbing sphere, a problem closely related to the volume of the Wiener sausage. In Fig. 2.12 I give a comparison of the expression given here, and the corresponding form based on Harris (1982). The two follow very similar trends, although the formulation given here better captures the diffusive limit. Given this similarity and the rule of parsimony, I will use Eq. 2.3.24 to describe the volume of the Wiener sausage.

The encounter rate of either an random walking predator (speed v, length scale λ) feeding on stationary prey (concentration C), or (by invariance) an ambush predator feeding on random walking prey (speed v, length scale λ, concentration C) is given by

$$Z = \frac{4C\lambda}{4\lambda + 3R}\pi R^2 v \tag{2.3.25}$$

These aspects of the Wiener sausage and encounter rates have important impact on the risk-benefit of particular search patterns as we will see in following sections.

3 Zooplankton Hydrodynamics

Interest in the hydrodynamics of zooplankton focuses on three main questions: how do zooplankton propel themselves through a fluid, how does this action of propulsion transmit information through hydromechanical signals to and from planktonic organisms, and how do the fluid flows produced by plankton influence the transport and dispersion of material dissolved or suspended in the fluid. From an ecological point of view, the question of propulsive mechanism is important in that it determines the efficiency of swimming. Specifically, since swimming speed is a determining factor in an organisms encounter rate with prey (and so energy in take), swimming efficiency figures in the cost-benefit balance of swimming behaviour. Perhaps more important is the action of moving through the fluid and the disturbance this generates. Many planktonic organisms are rheotactic – they can sense hydromechanical disturbances – predators can be alerted to the proximity of a prey organism, while the prey can be alerted to a predator. Fast swimming organisms may find more resources, but they are also hydromechanically more conspicuous and thus more prone to predator attacks than those that set a more leisurely pace. Stealth and early warning are twinned hydromechanical factors in the planktonic arms race, fine-tuned in the unforgiving arena of evolution by natural selection.

In order to give this topic its deserving rigour, I turn first to the nature of fluids, their basic characteristics (*e.g.* density, viscosity), and what the fluid environment means for a solid body, sinking or propelling itself through the fluid. What, for instance is the resistance of a fluid to a body moving through it, and how does the fluid deform in response to the body itself, and the forces applied by the body. These considerations lead from the Navier Stokes equations to the dynamics of a solid particle in a flow, encapsulated in the so-called Maxey-Riley-Gatingol equations. My chief concern here is low Reynolds number flow where many but not all zooplankton interactions take place. Along the way some mathematical abstractions are introduced – multipole expansions, Green's functions and spherical harmonics – tools which render at least some problems more tractable. Solutions and concepts found here give insight into both the problem of propulsion (not only those that are driven by the whips, hairs and paddles used by zooplankton but also other, more exotic mechanisms as well), and the patterns and potential information that can be exchanged between planktonic organisms. That moving organisms can be detected is reasonable enough, but what if they are totally immotile? As we shall see, there is no way for an organisms to remain 100% inconspicuous as its predator can deploy a hydromechanical radar.

3.1 Some basic hydrodynamics

3.1.1 The Navier-Stokes' equations

The Navier-Stokes' equations, named after Claude-Louis Navier[1] and George Gabriel Stokes[2], are a set of equations that describe the motion of fluids. They are essentially the fluid dynamic equivalent of Newtons 2nd law; force = mass × acceleration: that changes in momentum in infinitesimal volumes of fluid are simply the product of changes in pressure and viscous forces acting inside the fluid. Thus, the Navier-Stokes' equations are a dynamical statement of the balance of forces acting at any given region of the fluid.

$$\rho\frac{D\mathbf{u}}{Dt} = \rho\left(\frac{\partial\mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) = \mathbf{f} - \nabla p + \mu\nabla^2\mathbf{u} \tag{3.1.1}$$

Here, $\mathbf{u}$ is the fluid velocity, $\mathbf{f}$ is any external force (gravity for instance), p the pressure, μ the dynamic viscosity and ρ the density of the fluid. The force balance expressed above is for a fluid element; a hypothetical, infinitesimally small blob of fluid that moves, deforms and transmits fluid forces but remains identifiable at least for short periods of time. The plural "equations" is used because there are 3 of them; one each for each orthogonal direction.

Although not technically one of the Navier-Stokes equations, they are often written in conjunction with the equation of continuity (Daniel Bernoulli 1700-1782), which for an incompressible fluid is

$$\nabla\cdot\mathbf{u} = 0 \tag{3.1.2}$$

That is, since the fluid is incompressible, squeezing it somewhere means that it has to squirt out somewhere else. Sea water is very nearly incompressible: at 4000 m depth where the pressure is 400 times atmospheric pressure, its decrease in volume is only 1.8%.

Viscosity is a measure of the stickiness of a fluid – more specifically it says something about how rapidly a fluid deforms in response to given stresses. Consider for instance a sub-volume of a fluid bounded by a cube (Fig. 3.1). A stress (that is a force per unit area; F/A) applied on the surface of this cubic volume will cause it to deform. That is, vertical fluid lines within the volume will be stretched. This can be written

$$\frac{F}{A} = \mu\frac{1}{\xi}\frac{d\xi}{dt} = \mu\frac{1}{\ell}\frac{d\delta}{dt} = \mu\frac{\partial v}{\partial z} \tag{3.1.3}$$

where μ is the dynamic viscosity (dimensions $ML^{-1}T^{-1}$). The rate at which material lines of fluid are stretched (or compressed for that matter) is termed the rate of strain.

A fluid for which there is a linear relationship between stress and rate of strain is termed a Newtonian fluid. That is, a fluid that conforms to Newton's hypothesis as expressed in his *Philosophiae Naturalis Principia Mathematica*:

> " The resistance arising from imperfect slipping between fluid particles to be proportional to the velocity with which the particles are moving apart, all other things being equal"

[1] Claude-Louis Navier (1785–1836), a French physicist and engineer who contributed to the theory of elasticity and fluid mechanics.

[2] George Gabriel Stokes (1819–1903), a British mathematician and physicist. He held the chair of Lucasian Professor at Cambridge making contributions to the studies of light (*e.g.* diffraction and fluorescence) as well as fluid mechanics.

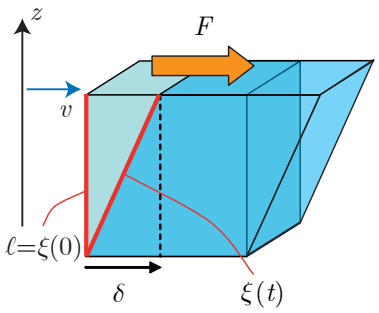


Figure 3.1: Viscosity and rate of strain for a Newtonian fluid (i.e. one that has no memory of shape). There are several equivalent ways of defining viscosity in terms of rate of strain. By noting that $d\xi/dt$, $d\delta/dt$ and v are all equivalent, Eq. (3.1.3) follows in each of its variations.

Often, viscosity is normalized by the fluid density ρ, $\eta = \mu/\rho$. This is termed the kinematic viscosity (dimensions L^2T^{-1}). The viscosity of seawater varies somewhat with temperature, and slightly dependent on salinity. A reasonable estimate for most cases is $\eta = 10^{-6}m^2s^{-1}$.

The force of gravity is often incorporated into the pressure term. If the fluid is at rest, the horizontal force balance requires that horizontal pressure gradients are zero. However, in the vertical direction (z measured positively upwards), the the vertical pressure gradient can be balanced by gravity:

$$0 = \rho g - \frac{\partial p_0}{\partial z} \tag{3.1.4}$$

where p_0 is the pressure in the absence of any motion. Integrating in the vertical this gives

$$p_0 = \int_0^z g\rho dz \tag{3.1.5}$$

When the fluid density is uniform (which is not always the case, but will suffice for the moment), this gives $p_0 = g\rho h$ where h is the depth below the surface. p_0 is termed the hydrostatic pressure.

A buoyant body - one that has a density less than its surrounding fluid - experiences an upward force. This of course is Archimedes' principle, or the law of upthrust:

> "Any body wholly or partially immersed in a fluid experiences an upthrust equal to, but opposite in sense to, the weight of the fluid displaced".

But what is this "upthrust" exactly? Returning to hydrostatic pressure, we see that integrating Eq. 3.1.4 over any given volume V within the fluid gives

$$g\int_V \rho dV = \int_V \frac{\partial p_0}{\partial z} dV = \int_S p_0 \hat{\mathbf{z}} \cdot d\mathbf{s} \tag{3.1.6}$$

That is, gravity acting on the mass within the volume is exactly balanced by the hydrostatic pressure acting over its surface. Introducing a body of different density ρ_b into V, does not alter the pressure acting over its surface. The net force f is thus given by

$$f = \int_V \frac{\partial p_0}{\partial z} dV - g\int_V \rho_b dV = g\int_V (\rho - \rho_b) dV \tag{3.1.7}$$

Thus, "upthrust" is the difference between gravity acting on the particle and hydrostatic pressure gradient – the latter being equal in magnitude but opposite in direction to gravity acting on the displaced fluid.

$D(\cdots)/Dt$ represents the substantive (or Lagrangian)[3] time derivative, expressing the rate of change following a fluid element. Conversely $\partial(\cdots)/\partial t$ represents the local (or Eulerian)[4] time derivative, expressing the rate of change at a fixed point in space. These are linked by an advection term $\mathbf{u}\cdot\nabla(\cdots)$, expressing the contribution of a local rate of change from the advection of gradients past the measuring point, or the measuring point moving through spatial gradients (Fig. 3.2). For a general function f,

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \mathbf{u}\cdot\nabla f \tag{3.1.8}$$

is the Euler-Lagrange transformation of rates of change. This can also be derived directly

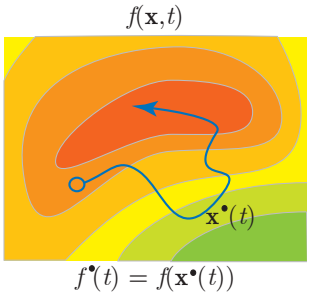


Figure 3.2: The Euler - Lagrange transformation. While this is often written in terms of the a differential operator (Eq. (3.1.8)) it actually arises from a much more fundamental identity; that a measurement should be the same no matter how observers move. This is expressed in Eq. (3.1.9).

from the Euler-Lagrange transformation of observed quantities. Namely that a quantity seen by an Eulerian observer must be identical to that seen by a Lagrangian observer at the instant when their positions coincide. Mathematically, this can be written

$$f^\bullet(t) = f\left(\mathbf{x}^\bullet(t)\right) \tag{3.1.9}$$

where $^\bullet$ indicates a Lagrangian measure, and $\mathbf{x}^\bullet(t)$ is the trajectory taken by the Lagrangian observer. Noting that velocity is the time rate of change of position, *i.e.* $\mathbf{u}(\mathbf{x}^\bullet, t) = d\mathbf{x}^\bullet/dt$, Eq. 3.1.8 follows from applying the chain rule of differentiation on Eq. 3.1.9.

The advection of momentum in Eq. 3.1.1, $\mathbf{u}\cdot\nabla\mathbf{u}$, expresses the fluid's inertia. The nonlinearities these introduce into the general Navier-Stokes dynamics make these equations not only difficult to solve, but are also the root of some important phenomena, in particular turbulence. There are occasions however, where these terms can be disregarded, greatly simplifying the mathematical treatment of fluid physics.

3.1.2 Relative motion near a point

The nature of a fluid is that it flows, and the rate at which it flows varies spatially. The kinematics of how this varies can be written in terms of the deformation and rotation of fluid elements in a small region of space. Specifically

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{x}_0) + \mathbf{e}\cdot(\mathbf{x}-\mathbf{x}_0) + \boldsymbol{\xi}\times(\mathbf{x}-\mathbf{x}_0) \tag{3.1.10}$$

[3] Joseph-Louis Lagrange, born Giuseppe Lodovico Lagrangia (1736 – 1813) was an Italian mathematician and astronomer, who lived most of his life in Prussia and France, making outstanding contributions to mathematics, and classical and celestial mechanics.

[4] Named after Leonhard Euler (1707 - 1783), a Swiss mathematician and physicist who made extensive contributions to the foundations of calculus and number theory introducing much of our modern terminology in mathematics (the concept of mathematical functions for instance). He also made extensive contributions to natural philosophy particularly in the fields of optics, mechanics, hydrodynamics and astronomy.

That is, at a location $\mathbf{x}$ not too distant from a reference point $\mathbf{x}_0$, the fluid velocity $\mathbf{u}$ is equal to that at $\mathbf{x}_0$ plus a contribution due to local deformation (rate of strain) and rotation. Essentially this is a first order Taylor expansion of the velocity field about $\mathbf{x}_0$. This can be formalized by the symmetric and anti-symmetric components of the local velocity gradients. Specifically, in Cartesian co-ordinates, we can define the two tensors:

$$e_{ij} = (u_{i,j} + u_{j,i})/2 \tag{3.1.11}$$

$$\xi_{ij} = (u_{i,j} - u_{j,i})/2 \tag{3.1.12}$$

(Batchelor, 1967).

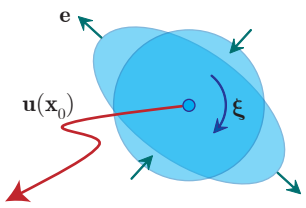


Figure 3.3: The relative motion of a fluid can be locally decomposed into translation of the point $\mathbf{u}(\mathbf{x}_0)$, the rate of strain $\mathbf{e}$ and rotation $\boldsymbol{\xi}$.

The matrices $\mathbf{e}$ and $\boldsymbol{\xi}$ have components e_{ij} and ξ_{ij} respectively and the notation $u_{i,j}$ implies $\partial u_i/\partial x_j$. The matrix $\mathbf{e}$, (or the tensor e_{ij}) represents the rate of strain of the fluid flow. This is also called the rate of deformation. The matrix $\mathbf{e}$ is symmetric about the diagonal; $e_{12} = e_{21}, e_{13} = e_{31}, e_{23} = e_{32}$. For an incompressible fluid, the divergence of the fluid flow (the trace of matrix $\mathbf{e}$) is zero. That is

$$e_{ii} = \nabla \cdot \mathbf{u} = e_{11} + e_{22} + e_{33} = 0 \tag{3.1.13}$$

The effect of $\mathbf{e}$ is most easily illustrated as how a spherical fluid volume is deformed into an ellipsoid (Fig. 3.3). A co-ordinate rotation can always be performed to realign these transverse components with some transformed co-ordinates. That is, the matrix $\mathbf{e}$ can be diagonalised. When in diagonal form, the elements along the diagonal are the principal components of the rate of strain. The feature of the flow field represented in the rate of strain tensor, is solely responsible for the motion of fluid elements relative to each other and thus the "resistance to imperfect slipping" – viscous effects – and the eventual dissipation of kinetic energy.

The matrix $\boldsymbol{\xi}$, (or the tensor ξ_{ij}) represents the rotation of the fluid flow, and may be written

$$\boldsymbol{\xi} = \frac{1}{2}\begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \tag{3.1.14}$$

where $\omega_n, n = 1, 2, 3$ represent the three components of vorticity. The anti-symmetric tensor represents solid body rotation, and thus, no relative motion of fluid elements.

3.1.3 Stress and strain in a fluid

We can now pass on from a consideration of the kinematics, to that of the dynamics of a fluid flow. A fluid element imbedded in a flow comes under the influence of several forces, some of which are body forces (*i.e.* intrinsic to the fluid element itself) and some of which are transmitted to the fluid element through its bounding surface from the surrounding fluid. To distinguish these, we can perform a thought experiment where we remove all the

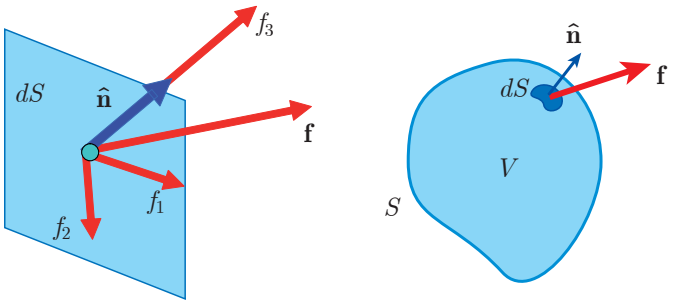


Figure 3.4: The force per unit area $\mathbf{f} = (f_1, f_2, f_3)$ acting on a small area dS with unit normal vector $\hat{\mathbf{n}}$. (In tensor notation, $f_i = \sigma_{ij}\hat{n}_j$ where σ_{ij} is the stress tensor). The total force acting on a small volume of fluid V with surface boundary S is the integral of stresses over its surface.

fluid surrounding the fluid element. The only force still active is gravity. This is a body force. (If the fluid element was electrically charged, and placed in a electric potential, as in electrophoresis, it would also come under the influence of an electrostatic body force.) Pressure and stresses only come into effect when the surrounding fluid is considered. These are forces that are extrinsic to the fluid element itself and act on the fluid element's boundary. Thus these effects are expressed as force per unit area.

If we consider a small area dS in a fluid, then the forces acting on it can be directed either perpendicular or parallel to its surface. In any particular orientation, the net force acting on the surface is given by

$$dF_i = f_i dS = \sigma_{ij}\hat{n}_j dS \tag{3.1.15}$$

where $\hat{\mathbf{n}} = \hat{n}_j$ is a unit vector normal to the surface, and $\boldsymbol{\sigma} = \sigma_{ij}$ is the stress tensor (Fig. 3.4).

For an incompressible, Newtonian fluid, the stress tensor can be written in terms of the pressure, p, the dynamic viscosity μ and the strain tensor $\mathbf{e} = e_{ij}$ as

$$\sigma_{ij} = -p\delta_{ij} + 2\mu e_{ij} \tag{3.1.16}$$

where δ_{ij} is the kronecker delta function: 1 for $i = j$ and 0 otherwise.

The force acting on a small volume of fluid V is thus the integral of f_i over the surface of volume. That is

$$F_i = \int_S f_i dS = \int_S \sigma_{ij}\hat{n}_j dS = \int_V \frac{\partial \sigma_{ij}}{\partial x_j} dV \tag{3.1.17}$$

where the latter step is an application of the divergence theorem – that the integral of divergence within a volume is equivalent to the flux integrated over its surface. That is

$$F_i = -V\frac{\partial p}{\partial x_i} + V2\mu\frac{\partial e_{ij}}{\partial x_j} \tag{3.1.18}$$

for an incompressible Newtonian fluid. Here I have used the property that for a very small volume V, the pressure gradient and strain tensor are essentially uniform within the volume. Thus we have

$$\frac{F_i}{V} = \frac{m}{V}\frac{Du_i}{Dt} = -\frac{\partial p}{\partial x_i} + 2\mu\frac{\partial e_{ij}}{\partial x_j} \tag{3.1.19}$$

where m is the mass of fluid within the volume. Of course m/V is the density of the fluid ρ. In this we have also used Newton's 2nd law: $F = mD\mathbf{u}/Dt$. Finally we can note that $e_{ij,j} = \frac{1}{2}\nabla^2\mathbf{u}$, so that the above, with the appropriate addition of a gravity body-force, is actually a derivation of the Navier-Stokes' equations.

3.1.4 Reynolds number


Osborne Reynolds

Examining the Navier-Stokes' equations (Eq. 3.1.1, Eq. 3.1.19), we can see that a fluid can either deform or accelerate in response to an applied force (a pressure gradient for instance). The Reynolds number says something about the relative effect a given force will have on these aspects of the fluid flow. In particular, suppose a flow has a characteristic flow speed U, and a characteristic size scale L. For a swimming organism for instance, U is its swimming speed and L is its body length. Without knowing too much in advance of the final flow, we can say that the flow will have a magnitude which is similar to U, and a spatial structure (*i.e.* velocity gradients) that are of size L. Thus, the acceleration term $\mathbf{u}\cdot\nabla\mathbf{u}$ (velocity times a velocity gradient) will have a size of the order U^2/L. On the other hand, the viscosity term $(\mu/\rho)\nabla^2\mathbf{u}$ will be of the order $(\mu/\rho)U/L^2$. The ratio of these is the Reynolds number

$$Re = \frac{\rho UL}{\mu} = \frac{UL}{\eta} \tag{3.1.20}$$

named after Osborne Reynolds[5]. When Re is much less than 1, viscous forces are large compared to inertia, and the term $\mathbf{u}\cdot\nabla\mathbf{u}$ can be ignored in the Navier-Stokes' equations. This case, where $Re << 1$ is known as the Stokes' regime. Conversely, if Re is much greater than 1, inertia dominates and the viscous term can be ignored. This latter case is known as the Euler regime.

3.1.5 The Stokes regime

Much of the drama of plankton in the marine environment is played out at low Reynolds numbers – the so called Stokes regime. In this world, the inertial terms can be taken to be negligible so that the Navier-Stokes' equations simplifies to

$$-\nabla p + \mu\nabla^2\mathbf{u} = \mathbf{f} \qquad ; \qquad \nabla\cdot\mathbf{u} = 0 \tag{3.1.21}$$

These, together with the equation of continuity for an incompressible fluid are often referred to as the Stokes' equations as they define the fundamental physics of the Stokes' regime. Importantly, these equations are linear. That is, the solution for an external force $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, is simply the sum of the solution for $\mathbf{f}_1$ and $\mathbf{f}_2$.

In many cases, the external force is zero, and the flow is prescribed by boundary conditions. A rigid body for instance moving through the fluid requires that 1) there is no flow perpendicular to the body's surface (fluid cannot flow in or out of the body), and given that the fluid is viscous 2) that there is no slip between the fluid and the body. What this means is that if the solid body is moving with velocity $\mathbf{v}$, every fluid element on the surface of the body (*i.e.* at $\mathbf{x} = \mathbf{s}$) must also be moving with velocity $\mathbf{v}$. The flow is then prescribed by a boundary

[5]Osborne Reynolds (1842-1912), a British physicist most famous for his insights into the transition from laminar flow to turbulence; a transition described my the Reynolds number.

value problem

$$-\nabla p + \mu\nabla^2\mathbf{u} = 0 \quad \text{subject to} \quad \begin{cases} \mathbf{u}(\mathbf{x} = \mathbf{s}) = \mathbf{v} \\ \mathbf{u}(\mathbf{x} \to \infty) = 0 \end{cases} \tag{3.1.22}$$

The second condition listed here states that whatever the effect of the moving body on the fluid, the flow it induces must disappear far away. It is also possible to change the reference system from the laboratory frame (*i.e.* within which the fluid at infinity is motionless) to one fixed on a moving particle. This

$$-\nabla p + \mu\nabla^2\mathbf{u} = 0 \quad \text{subject to} \quad \begin{cases} \mathbf{u}(\mathbf{x} = \mathbf{s}) = 0 \\ \mathbf{u}(\mathbf{x} \to \infty) = -\mathbf{v} \end{cases} \tag{3.1.23}$$

Here, the particle is fixed, and the fluid moves with respect to the observer.

One final general property of the Stokes' regime is that by whatever means motion is imposed on the fluid, its response is instantaneous. That is, there is no inherent time response of the fluid, and any time variation seen in the system is due solely to time variations in the applied forces or boundary conditions. This has immediate implications for swimming organisms since it leads to rate independence and the requirement of time reversibility: that the swimming action must look different when "played" (as in a video recording) forwards or backwards in order for net propulsion to be achieved at low Reynolds numbers.

3.1.6 The Euler regime

At the other extreme, when Reynold's numbers are high ($Re \gg 1$), viscous forces become negligible, and the Navier-Stokes' equation at steady state simplifies to:

$$\nabla p + \mathbf{u} \cdot \nabla\mathbf{u} = \mathbf{f} \tag{3.1.24}$$

In general, the Euler regime is much difficult analyze than the Stokes' regime, there being complicating non-linear terms $\mathbf{u} \cdot \nabla\mathbf{u}$ involved. There are however accessible solutions – potential flows – that are relevant for plankton hydrodynamics.

Potential flow: A flow is said to be irrotational if its vorticity is zero. That is

$$\boldsymbol{\omega} = \nabla \times \mathbf{u} = 0 \tag{3.1.25}$$

For a body moving through a fluid, vorticity is generated by shear stresses between the body's surface and the fluid. For an inviscid fluid (*i.e.* in the Euler regime) no vorticity is generated by this mechanism, so if a flow is irrotational, it remains irrotational even though some object moves through it. An elementary identity in vector calculus states that $\nabla \times \nabla\phi = 0$ where ϕ is any scalar field, so that an irrotational flow field can be expressed as the gradient of a scalar field the scalar potential – and $\mathbf{u} = \nabla\phi$. Such a flow field is said to be lamellar (not to be confused with laminar). A corollary to this is a non-divergent flow where $\nabla \cdot \mathbf{u} = 0$ suggesting that the flow field can be written in terms of a vector potential $\mathbf{q}$ such that $\mathbf{u} = \nabla \times \mathbf{q}$. A non-divergent flow field is termed solenoidal. If a flow field is both irrotational and non-divergent, then it is a Laplacian field, subject to the constraint $\nabla^2\phi = 0$.

As an example, the potential flow around a sphere of radius a moving at speed U through and *inviscid* fluid can be written in terms of the stream function

$$\psi = -U\left(\frac{r^2}{2} - \frac{a^3}{2r}\right)\sin^2(\theta) \tag{3.1.26}$$

from which the scalar potential

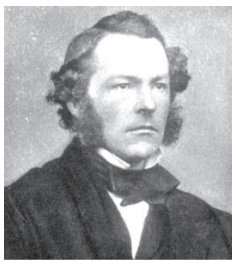
$$\phi = -\left(r - \frac{a^3}{2r^2}\right)\cos(\theta) \tag{3.1.27}$$

can be derived.

3.2 Solid bodies in fluids

Plankton and detrital material are first and foremost solid bodies imbedded in a fluid. Their movement through the fluid is due either to gravity, or because they themselves undulate or move appendages.

3.2.1 Creeping flow around a sphere


George Gabriel Stokes

In the Stokes regime, the equations of motion for a fluid are linear. For a large class of problems for a fluid interacting with a solid, rigid particle, this is simplified even further by noting that some physical settings have an axial symmetry; for instance flows interacting with a rigid spherical particle. Under these conditions, the flow field is a function of the radial (r) and zenith or polar (θ) coordinates, but independent of the azimuthal (ϕ) where $\theta = 0$ defines the axis of symmetry (*cf.* Fig. 3.5). The flow can be characterized by a stream function ψ, which describes the trajectory of fluid elements on a (r, θ) plane. The velocity field in spherical polar coordinates then is given by

$$u_r = \frac{1}{r^2 \sin\theta}\frac{\partial\psi}{\partial\theta} \tag{3.2.1}$$

$$u_\theta = -\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r} \tag{3.2.2}$$

It may not be immediately obvious, but these forms of (u_r, u_θ) comply with the non-divergence condition ($\nabla \cdot \mathbf{u} = 0$) for spherical polar coordinates with $u_\phi = 0$, and $\partial_\phi = 0$). A convenient way of expressing the velocity field is

$$\mathbf{u} = \nabla \times \left(\hat{\boldsymbol{\phi}}\frac{\psi}{r\sin\theta}\right) \tag{3.2.3}$$

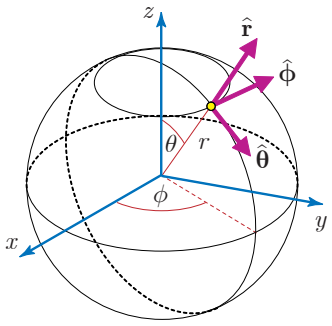


Figure 3.5: Spherical polar coordinates: in many cases, the solution to a problem is facilitated by its geometry. Spherical polar coordinates in the radial (r), zenith (θ) and azimuthal (ϕ) directions can be related to cartesian coordinates (x, y, z) as $r^2 = x^2 + y^2 + z^2$, $\tan\phi = y/x$ and $\tan\theta = \sqrt{x^2 + y^2}/z$. Alternatively, cartesian coordinates can be related to spherical polar coordinates as $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\sin\phi$.

The Navier-Stokes' equations, together with the non-divergence property of an incompressible flow can be used to give the following properties

$$0 = \nabla^2 p \tag{3.2.4}$$
$$0 = \nabla^2(\nabla \times \mathbf{u}) \tag{3.2.5}$$

The first comes from the divergence of the Navier-Stokes' equations, and the second from their curl. These are simplified respectively by observing that $\nabla \cdot \mathbf{u} = 0$ and $\nabla \times \nabla p = 0$. It is convenient to write $\nabla \times \mathbf{u} = \boldsymbol{\omega}$, which can be expanded in terms of the stream function as

$$\begin{aligned} \boldsymbol{\omega} &= \nabla \times \mathbf{u} = \nabla \times \nabla \times \left(\hat{\boldsymbol{\phi}} \frac{\psi}{r \sin\theta} \right) \\ &= -\hat{\boldsymbol{\phi}} \frac{1}{r \sin\theta} \left(\frac{\partial^2 \psi}{\partial r^2} + \frac{\sin\theta}{r^2} \frac{\partial}{\partial \theta} \left(\frac{1}{\sin\theta} \frac{\partial \psi}{\partial \theta} \right) \right) \end{aligned} \tag{3.2.6}$$

Some useful formulae with regards to this derivation can be found in Batchelor (1967). $\boldsymbol{\omega}$ is the vorticity of the flow field; it is a vector in the azimuthal direction $\hat{\boldsymbol{\phi}}$. One additional useful vector identity is that for an incompressible fluid, the Laplacian of the flow field is equal and opposite to the curl of the vorticity. That is

$$\nabla^2 \mathbf{u} = -\nabla \times \nabla \times \mathbf{u} = -\nabla \times \boldsymbol{\omega} \tag{3.2.7}$$

which, when combined with Eq. 3.2.5 and Eq. 3.2.6 leads to

$$\left(\frac{\partial^2}{\partial r^2} + \frac{\sin\theta}{r^2} \frac{\partial}{\partial \theta} \left(\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \right) \right)^2 \psi = 0 \tag{3.2.8}$$

This is a general expression for a stream function for any axially symmetric creeping flow expressed in spherical polar coordinates. Any specific application is subject not only to the dynamics expressed in Eq. 3.2.8 but also to boundary conditions; how the flow behaves on the surface of solid bodies and at large distances away from applied forces.

As an important example, consider a sphere of radius a moving through a fluid with velocity $\mathbf{U}$. The defining boundary conditions are that (1) the flow $\mathbf{u} = 0$ on the surface of the sphere $r = a$, and that at a large distance from the sphere $r \to \infty$, the flow relative to the sphere becomes equal to its translational velocity, $\mathbf{u} = \mathbf{U}$. Aligning the axis of symmetry with the direction of translation, the the stream function is

$$\psi \to -\frac{U}{2} r^2 \sin^2\theta \tag{3.2.9}$$

for $r \to \infty$. The form of this asymptotic solution suggests a trial solution of the form

$$\psi(r, \theta) = f(r) \sin^2\theta \tag{3.2.10}$$

Substituting into Eq. 3.2.8 this gives

$$\left(\frac{\partial^2}{\partial r^2} - \frac{2}{r^2} \right)^2 f = 0 \tag{3.2.11}$$

Thus the general solution should look like

$$\psi = -\frac{U}{2} \left(r^2 + c_1 a r + c_2 \frac{a^3}{r} \right) \sin^2\theta \tag{3.2.12}$$

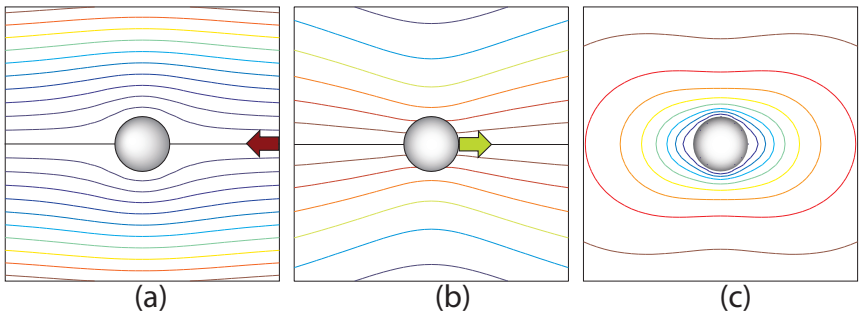


Figure 3.6: Stream functions for flow around a sphere moving through a viscous fluid; (a) relative to the sphere, (b) relative to the far field flow, and (c) the isotachs (contours of relative speed between the sphere and the fluid) in intervals of 0.1 times the far field velocity difference.

The boundary condition that $u_r = 0$ at $r = a$ means that $1 + c_1 + c_2 = 0$, and $u_\theta = 0$ at $r = a$ requires that $2 + c_1 - c_2 = 0$. Thus, $c_1 = -3/2$ and $c_2 = 1/2$, and the full solution is

$$\psi = -\frac{U}{2}\left(r^2 - \frac{3}{2}ar + \frac{1}{2}\frac{a^3}{r}\right)\sin^2\theta \tag{3.2.13}$$

(Fig. 3.6). Given this form of the stream function, it is relatively straight forward to determine the velocity components. Substituting Eq. (3.2.13) into Eq. (3.2.1) and Eq. (3.2.2) gives

$$u_r = U\left(\frac{1}{2}\left(\frac{a}{r}\right)^3 - \frac{3}{2}\left(\frac{a}{r}\right) + 1\right)\cos\theta \tag{3.2.14}$$

$$u_\theta = U\left(\frac{1}{4}\left(\frac{a}{r}\right)^3 + \frac{3}{4}\left(\frac{a}{r}\right) - 1\right)\sin\theta \tag{3.2.15}$$

From the Navier-Stokes' equations in the Stokes limit, this gives the pressure gradient components as:

$$\frac{\partial p}{\partial r} = \left(\frac{3\mu a U}{r^3}\right)\cos\theta \tag{3.2.16}$$

$$\frac{\partial p}{\partial \theta} = \left(\frac{3\mu a U}{2r^2}\right)\sin\theta \tag{3.2.17}$$

These can be solved to give

$$p = p_\infty - \left(\frac{3\mu a U}{2r^2}\right)\cos\theta \tag{3.2.18}$$

where p_∞ is the non-spatially varying component of the pressure.

In general, the stress components in the flow ate given by:

$$\sigma_{rr} = -p + 2\mu e_{rr} = -p + 2\mu\frac{\partial u_r}{\partial r} \tag{3.2.19}$$

$$\sigma_{r\theta} = 2\mu e_{r\theta} = \mu r\frac{\partial}{\partial r}\left(\frac{u_\theta}{r}\right) + \frac{\mu}{r}\frac{\partial u_r}{\partial \theta} \tag{3.2.20}$$

The stress components on the surface of the sphere $r = a$ are

$$\sigma_{rr}(a,\theta) = -p_\infty + \left(\frac{3\mu U}{2a}\right)\cos\theta \tag{3.2.21}$$

$$\sigma_{r\theta}(a,\theta) = -\left(\frac{3\mu U}{2a}\right)\sin\theta \tag{3.2.22}$$

from which the net force per unit area in the z (*i.e.* the direction of motion) direction can be specified:

$$f_z(a,\theta) = \sigma_{rr}(a,\theta)\cos\theta - \sigma_{r\theta}(a,\theta)\sin\theta = -p_\infty + \frac{3\mu U}{2a}$$

Integrating over the surface of the sphere gives

$$F = 2\pi a\int_0^\pi f_z(a,\theta)a\sin\theta d\theta = 6\pi\mu U a \tag{3.2.23}$$

which is the net drag on the sphere. This is Stokes' law, and relates the resistance active on a sphere moving through a fluid to its speed, size and the fluid's viscosity.

3.2.2 The boy and his balloon

The allegory of small scale fluid dynamics being non-intuitive to every day human experience is usually couched in terms of viscosity – that a small organism swimming in water is analogous to us trying to propel our-selves through syrup. There is however a much more fundamental aspect of fluid dynamics that I've found can confound even experience physicists – one that has to do with buoyancy, inertia and pressure gradients.

The basic tenets of the problem can be illustrated in a simple thought experiment. A boy sits in an aeroplane holding a helium filled balloon by a string. While stationary at the end of the runway before takeoff, the balloon being buoyant keeps the string perfectly vertical. As the aeroplane accelerates down the runway the boy and everybody else is pressed backwards into their seats, but which way does the balloon move - towards the front of the plane, towards the back of the plane, or does it remain vertical?

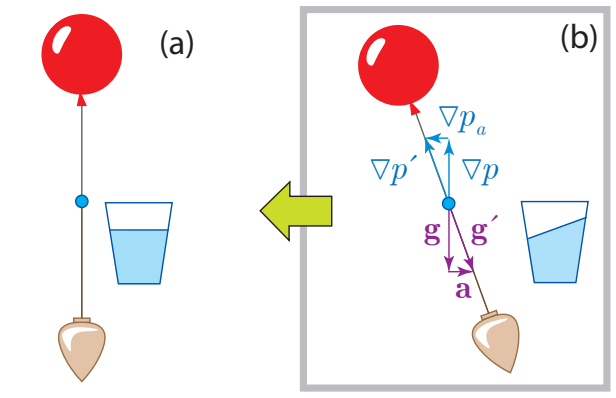


Figure 3.7: A balloon, plumb-bob, and a glass of water in an air filled space (a) under gravity alone, and (b) with an additional horizontal acceleration.

The correct answer is that the balloon moves forward. A somewhat glib proof of this comes from appealing to Einstein's theory of general relativity. Specifically, a man in a box in

outer space cannot tell the difference between the acceleration of the box, or the presence of a gravitational field. The laws of physics should be identical. Thus, for the boy and his balloon, as the plane accelerates there is an additional component to the effective gravity, $\mathbf{g}' = \mathbf{g} + \mathbf{a}$ with the $\mathbf{g}$ the rest frame gravitational acceleration down, and $\mathbf{a}$ the acceleration of the plane directed backwards. For the accelerating plane, "down" is directed somewhat to the rear of the plane, while "up" is tilted forward. Thus, for the laws of physics to be consistent, the balloon remains aligned with the new "up" direction and so tilts towards the front of the plane (Fig. 3.7).

This is not really satisfactory as it is only a phenomenological explanation, appealing only so far as Archimedes' principle of upthrust. In more detail what happens, is that in accelerating, the air enclosed within the plane establishes a pressure gradient in the direction of acceleration. This is why the air remains at rest relative to the boy and his balloon, in a manner directly analogous to hydrostatic pressure. This pressure gradient acts on all bodies in the plane, but for most, their density being much greater than air, this force is almost negligible. For the balloon, its density being less than air, this force becomes significant and the balloon moves forward.

Near neutrally buoyant bodies immersed in a fluid experience pressure gradient forces that we have no analogue for. For plankton, near neutral buoyancy is a defining property, and their motion pretty much follows the motion of the fluid – the forces active on a volume of fluid are very nearly equal to the forces that are active on a solid body occupying the same volume.

3.2.3 Motion of a sphere in unsteady flow

How is the movement of individual plankters affected by the motion of a fluid, and how, in turn does the fluid motion respond to the movement of individual plankters. These are fundamental questions that can be addressed, at least in an idealized form, through examining the hydrodynamics of small rigid spherical particles at low Reynolds number. Plankton, of course are not spherical, neither are they rigid. None-the-less, these assumptions are a reasonable point to start from. The equation of motion for such a particle (radius a, mass m_p) in a fluid is given by

$$\begin{aligned} m_p \frac{d\mathbf{v}}{dt} =& m\frac{D\mathbf{u}}{Dt} + (m_p - m)\mathbf{g} \\ & - 6\pi a\mu \left(\mathbf{v} - \mathbf{u} - \frac{a^2}{6}\nabla^2\mathbf{u}\right) \\ & - \frac{m}{2}\left(\frac{d\mathbf{v}}{dt} - \frac{D}{Dt}\left(\mathbf{u} + \frac{a^2}{10}\nabla^2\mathbf{u}\right)\right) \\ & - 6\pi a^2\mu \int_0^t dt' \left(\frac{1}{\sqrt{\pi\eta(t-t')}}\left(\frac{d}{dt'}\left(\mathbf{v} - \mathbf{u} - \frac{a^2}{6}\nabla^2\mathbf{u}\right)\right)\right) \end{aligned} \tag{3.2.24}$$

where $\mathbf{v}$ is the velocity of the particle, $\mathbf{u}$ is the undisturbed velocity of the fluid, $\mathbf{g}$ is gravitational acceleration (direction vertically down) and m is the mass of displaced fluid ($m = 4\pi a^3\rho/3$). The dynamic viscosity μ is related to the kinematic viscosity η by $\mu = \rho\eta$ where ρ is the density of the fluid (Maxey and Riley, 1983; Gatignol, 1983).

The derivation of these equations, initiated by Stokes in the 1840s, has developed over the years, leading to ever more refinements. A relatively full accounting of the physics involved

can be found in Michaelides (1997) a treatment that is relatively easy to understand as well. Often, these equations are termed the Maxey-Riley-Gatingol equations[6].

The time derivative operators d/dt and D/Dt are both Lagrangian, but have a different meaning; the former d/dt denotes the time rate of change following the spherical particle itself whereas D/Dt denotes the time rate of change following a fluid element in the undisturbed flow field, evaluated at the particle location. Formally this is expressed as:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \tag{3.2.25}$$

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla = \frac{d}{dt} + (\mathbf{u} - \mathbf{v}) \cdot \nabla \tag{3.2.26}$$

The terms of the right hand side of Eq. (3.2.24) represent

- The force acting on the "undisturbed" fluid (*i.e.* as if the particle wasn't there) which is now active on the particle. This is composed of both a rotational contribution and a contribution due to translation. The latter is of particular importance in many applications, and arises from the mean pressure gradient force acting over the surface of the sphere – much like the boy and his balloon.

- The buoyancy force; $\mathbf{g}$ is the acceleration due to gravity and is directed vertically down. This is a balance between the downward force of gravity on the particle, and the upward force of the hydrostatic pressure gradient acting on the fluid. The principle involved here has been known since antiquity.

- Drag. This is essentially due to viscous forces and velocity differences between the particle and the fluid. This is sometimes referred to a Stokes' drag (in deference to the low Reynolds' number regime), and includes the skin drag and form (pressure) drag. The spherical shape of the particle is assumed in this term, a calculation first introduced by G.G. Stokes in 1851.

- The added mass force. Added mass is the inertia added to a system due to the fact that an accelerating body must also accelerate some volume of surrounding fluid as it moves through it, since the object and fluid cannot occupy the same physical space simultaneously. For a sphere the added mass is $2\pi a^3\rho/3$ (curiously derived by integrating the inertia associated with a potential *i.e.* and inviscid flow). The added mass is half the mass of the displaced fluid. The concept of added mass was introduced in 1828 by the German mathematician Frederich Bessel (1784–1846), and subsequently re-visited by Simèon Denis Poisson (1831) and George Green (1833). All of these endeavours were initiated, at least in part, in connection with precise calculation on the period of a pendulum, accurate time-keeping, and the question of the determination of longitude at sea.

- The Basset-Boussinesq history force. There is a considerable volume of fluid adjacent to the moving body that is dragged along due to viscosity. This water mass has associated inertia that has to be accelerated in order for the particle to speed up or slow down. It is this inertial effect that is represented in the history force.

[6]These are also refereed to as the Maxey-Riley equations, but recent usage prefers that given above in recognition of Gatingol's independent derivation.

- Terms in $a^2\nabla^2\mathbf{u}$ are the Faxén correction terms and are related to the curvature of the flow field. Simply put, when the undisturbed flow field $\mathbf{u}$ varies significantly over the spatial dimensions of the particle, the undisturbed velocity at the centre of the particle may differ from the average undisturbed fluid flow over the volume of the particle. The latter is more pertinent with respect to the particle motion as this would more directly reflect the mean pressure gradient and drag applied to the particle, a point made by Faxén in his Ph.D. thesis in 1921.

Eq. (3.2.24) is valid provided:

$$a(\mathbf{v}-\mathbf{u})/\eta \ll 1 \tag{3.2.27}$$
$$a/L \ll 1 \tag{3.2.28}$$
$$a^2U/(\eta L) \ll 1 \tag{3.2.29}$$

The first of these expresses the low Reynolds number criterion. The second stipulates that the size of the particle should be considerably less than the length scale over which the fluid flow varies, while the third states that the Reynolds number associated with the velocity difference across the particle (the gradient Reynolds number) must be much less than 1. These conditions assure that (1) the drag between the fluid and the particle is a linear function of the velocity difference, (2) that the Faxén correction terms for curvature are valid to leading order, and (3) that the perturbation to the flow field introduced by the presence of the solid particle is small and is well approximated by the added mass and history terms. The full Maxey-Riley-Gatingol equations are rather cumbersome in their integro-differential form, and in many, if not all applications, the equations are simplified in one way or the other.

3.2.4 Physical background to the Maxey-Riley-Gatingol equations

In a very general sense, the force $\mathbf{F}$ acting on a moving solid body in an unsteady flow field can be divided into 3 contributions:

$$\mathbf{F} = \mathbf{F}_f + \mathbf{F}_b + \mathbf{F}_i = m_p\frac{d\mathbf{v}}{dt} \tag{3.2.30}$$

where $\mathbf{F}_f$ is the force induced by the undisturbed flow, $\mathbf{F}_b$ is the body force (*e.g.* gravity) acting on the solid body, and $\mathbf{F}_i$ is the sum of forces that arise through interactions of the flow and the solid body. If we write Ω as representing the volume of space occupied by the solid body, and the density of the fluid is uniform, the first 2 terms in this force balance can be readily deduced to be the sum of the volume averaged fluid acceleration in Ω, hydrostatic pressure, and gravity acting on the solid body proper. That is

$$\mathbf{F}_f + \mathbf{F}_b = m\left\langle\frac{D\mathbf{u}}{Dt}\right\rangle_\Omega + (m_p - m)\mathbf{g} \tag{3.2.31}$$

where $\langle\cdots\rangle_\Omega$ indicates a volume average, and is true for any size or shape of particle.

The force $\mathbf{F}_i$ induced by the flow perturbation however is the real test, and where most of the approximations and subsequent restrictions to the Maxey-Riley-Gatingol equations have their source. The perturbation flow force arises from two physical processes expressed in 3 terms; that a force accelerating the particle must also accelerate a volume of surrounding

fluid (added mass force $\mathbf{F}_a$), and the fluid exerts a viscous drag on the body due to any velocity differences between its surface and the surrounding fluid (drag force $\mathbf{F}_d$). Finally, the time evolution of this viscous boundary layer also involves fluid inertia and appears as a third term (history force $\mathbf{F}_h$). In summary

$$\mathbf{F}_i = \mathbf{F}_a + \mathbf{F}_d + \mathbf{F}_h \tag{3.2.32}$$

The effect of added mass can be most clearly seen when all other influences (gravity and viscosity) are "turned off". In this case, the force on a solid body imbedded in a uniform flow field can be written:

$$\mathbf{F}' = \rho\mathcal{V}\left\{(1+C_m)\frac{D\mathbf{u}}{Dt} - C_m\frac{d\mathbf{v}}{dt}\right\} \tag{3.2.33}$$

where $\mathcal{V}$ is the particle volume, ρ is the fluid density and C_m is the added mass coefficient; the apparent volume of fluid co-accelerated with the particle expressed as a fraction of the particle volume (Auton et al., 1988). One of the earliest results in fluid particles interactions was by Poisson (1831), where he showed that the added mass coefficient for a sphere was $C_m = \frac{1}{2}$. Soon after, Green (1833) demonstrated that the added mass coefficient for an ellipsoid was also $C_m = \frac{1}{2}$. That is, for a force applied directly to a particle, its rate of change of momentum will be reduced by a given amount equivalent to the rate of change of momentum of one half its volume of ambient fluid exposed to the same force. For a non-uniform flow field, the evaluation requires a volume integral where we recognize the first term as being the force exerted by the undisturbed flow. The remaining contribution is the added mass force, and for an ellipsoid is written as:

$$\mathbf{F}_a = \frac{m}{2}\left\{\left\langle\frac{D\mathbf{u}}{Dt}\right\rangle_\Omega - \frac{d\mathbf{v}}{dt}\right\} \tag{3.2.34}$$

The drag experienced by a sphere moving with a uniform speed relative to a viscous fluid was one of the first practical applications of the Navier-Stokes' theory (Stokes, 1845, 1851); $\mathbf{F}'_d = 6\pi\mu a\mathbf{U}$, where $\mathbf{U}$ is the relative speed, provided $Re = a|\mathbf{U}|/\eta \ll 1$. It is a fairly simple matter to extend this result for the drag experienced by a spherical particle in an unsteady, nonuniform flow:

$$\mathbf{F}_d = 6\pi\mu a\left\{\langle\mathbf{u}\rangle_{d\Omega} - \mathbf{v}\right\} \tag{3.2.35}$$

for the condition $Re = a|\langle\mathbf{u}\rangle_{d\Omega} - \mathbf{v}|/\eta \ll 1$. It is in formulating this component that most of the restrictions on the Maxey-Riley-Gatingol equations arise. Before proceeding with these restrictions, the viscous interaction between the particle and the fluid leads to an additional dynamic effect: the Basset-Boussinesq history force. This arises due to time variations in the drag force exerted by the particle on the fluid. There is a finite time scale of the order a^2/η for the transmission of these variations into changes in the fluid's inertia in the viscous boundary layer. The first to derive an expression for this was Boussinesq (1885) followed independently and soon after by Basset (1888); hence the nomenclature. Without delving into the derivation specifically, the history term can be written:

$$\mathbf{F}_h = \frac{a}{\sqrt{\pi\eta}}\int_{-\infty}^{t}\frac{d\mathbf{F}_d}{ds}\frac{ds}{\sqrt{t-s}} \tag{3.2.36}$$

Note that this will always have a buffering effect of the particle motion compared to the fluid.

Appearing in the added mass, drag, and by extension the history terms are contributions of the form $\langle \mathbf{u} \rangle_\Omega$ and $\langle \mathbf{u} \rangle_{d\Omega}$; *i.e.* the average of the fluid flow velocity over the volume and surface of the particle respectively. The point to note is that in general, the undisturbed flow field is non-uniform and contains spatial gradients of velocity. Thus, estimates of the surface and volume averages may differ somewhat from the evaluation of the undisturbed velocity at the centre of mass of the solid body. Provided these gradients are not too large, and the solid body is assumed to be spherical, the evaluation of the averages can be achieved by a Taylor expansion of the velocity field around the centre of mass. Specifically

$$\left\langle \mathbf{u}^{(0)} \right\rangle_{d\Omega} = \mathbf{u}^{(0)}(\mathbf{x}_0, t) + \frac{a^2}{6}\nabla^2 \mathbf{u}^{(0)}(\mathbf{x}_0, t) + O(a^4) \tag{3.2.37}$$

$$\left\langle \mathbf{u}^{(0)} \right\rangle_{\Omega} = \mathbf{u}^{(0)}(\mathbf{x}_0, t) + \frac{a^2}{10}\nabla^2 \mathbf{u}^{(0)}(\mathbf{x}_0, t) + O(a^4) \tag{3.2.38}$$

This is the origin of the Faxén correction terms for flow curvature. Such an approach is valid only if the size of the particle is small compared to the size of structures within the flow field: $a/L \ll 1$.

The final point to note is that in deriving this formulation, the focus has been on how the motion of the solid body can be deduced from the undisturbed fluid flow (denoted $\mathbf{u}^{(0)}$). The added mass, drag and history terms are all manifestations of this, and each arising due to an apparent "slip" velocity between the solid body and the undisturbed flow field ($\mathbf{w} = \mathbf{v} - \mathbf{u}^{(0)}$). Kinematically however, no such "slip" velocity can exist as the boundary conditions that on the surface of the moving body must obey $\mathbf{u}_{d\Omega} = \mathbf{v}$. That is, there must be a perturbation flow field (denoted $\mathbf{u}^{(1)}$), generated by the interaction of the undisturbed flow and the solid body that ensure the non-slip boundary condition on the surface of the solid body. Thus, without delving too deeply into the dynamics of the system, it can be readily estimated that $O(\mathbf{u}^{(1)}) = O(\mathbf{w})$, and in order for the perturbation velocity to conform to a low Reynolds number, the additional criterion that $a^2 U/\eta L \ll 1$.

The proper inclusion of the perturbation flow in the dynamic analysis has immediate consequences for inertial effects, specifically those arising through the distinction between the fluid acceleration (D/Dt) and the particle acceleration (d/dt). While conceptually important, the inclusion of this distinction in the Maxey-Riley-Gatingol equations is largely cosmetic in that the neglected flow perturbation term will be of similar magnitude as the purported correction:

$$D/Dt - d/dt = \mathbf{w} \cdot \nabla \approx \mathbf{u}^{(1)} \cdot \nabla$$

This point made by Maxey and Riley (1983) in their original derivation is often ignored.

3.2.5 Simplified Maxey-Riley-Gatingol equations

By ignoring the Basset-Boussinesq history term and the Faxén correction terms (which are generally all of small order) Eq. (3.2.24) can be rewritten in a simple form

$$\frac{d\mathbf{v}}{dt} = \beta \frac{D\mathbf{u}}{Dt} + (1-\beta)\mathbf{g} - \frac{1}{\tau}(\mathbf{v} - \mathbf{u}) \tag{3.2.39}$$

where $\beta = 3m/(2m_p + m)$ and the Stokes' time scale $\tau = \rho a^2/(3\beta\mu) = a^2/(3\beta\eta)$. For a heavy particle ($\rho_p \gg \rho$), $\beta \to \rho/\rho_p \to 0$, while for a light particle ($\rho_p \ll \rho$), $\beta \to 3$. For a neutrally buoyant particle ($\rho_p = \rho$), $\beta = 1$.

One parameter that is of immediate relevance to the following discussion is the Stokes' number $St = \tau/T$, defined as the ratio of the Stokes' time scale τ to the time scale associated with the fluid acceleration T. It says something about the time associated with the viscous adjustment of velocity differences between the particle and the fluid. Importantly, the Stokes' number comes in 2 flavours – one due to inherent time variations in the local flow, and the other due to fluid elements being advected through velocity gradients in a steady flow field. While the former places no restrictions on the validity of the Maxey-Riley-Gatingol equations, the latter does. Specifically, if U is the magnitude of the flow velocity, and L is the length scale of spatial gradients, then $T \sim L/U$ so that

$$St_\nabla = \frac{U}{L}\frac{a^2}{3\eta} \tag{3.2.40}$$

where St_∇ is used to distinguish this Stokes' number as being associated with gradients. We recognise this number as being involved in one of the criteria (Eq. 3.2.29) under which the derivation of the Maxey-Riley-Gatingol equations is valid. Specifically

$$St_\nabla = \frac{U}{L}\frac{a^2}{3\eta} \ll 1 \tag{3.2.41}$$

That is, the gradient Stokes' number is essentially the same as the gradient Reynolds number, and both must be small.

Sinking velocity: An immediate application for this formula is in calculating the sinking speed of spherical bodies in a quiescent fluid. Specifically, assuming the $\mathbf{u} = 0$ (the fluid is motionless), then

$$\frac{d\mathbf{v}}{dt} = (1-\beta)\mathbf{g} - \frac{1}{\tau}(\mathbf{v}-\mathbf{u}) \tag{3.2.42}$$

At steady state ($\mathbf{v}$ = constant), a force balance between buoyancy and friction gives the settling speed of

$$\mathbf{v} = \mathbf{g}\frac{2}{9}\frac{a^2}{\eta}\frac{\rho_p - \rho}{\rho} \tag{3.2.43}$$

For a particle denser than the fluid ($\rho_p > \rho$), this velocity is in the same direction as gravity – logically enough it sinks. (Fig. 3.8). For $\rho_p < \rho$, the particle is buoyant and rises. For an

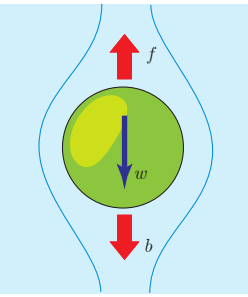


Figure 3.8: Settling velocity. The same result follows from Stokes law; that the viscous force on a spherical particle (radius a) moving with speed v through a fluid (dynamic viscosity μ) is $F_\mathrm{d} = 6\pi a\mu v$. This can be equated with the net gravitational force (i.e. gravity down - buoyant uptrust) on the same spherical particle, $F_\mathrm{g} = 4g\pi a^3(\rho_p - \rho)/3$. Equating these two and solving for v gives Eq. (3.2.43).

air bubble in water, $\rho \approx 1000\rho_p$ so that $\mathbf{v} = -\mathbf{g}2/9a^2/\eta$.

Simplified Maxey-Riley-Gatingol equations II, Heavy approximation: An even greater simplification can be made when the density of the particle is much greater than the carrier fluid.

In this case $\beta \to 0$ so that

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{\tau}(\mathbf{v} - \mathbf{u}) \tag{3.2.44}$$

and $\tau \to 2a^2\rho_p/(9\mu\rho)$. In a natural setting, this approximation is mostly valid when the carrier fluid is air as for aerosols and water droplets ($\rho_p \approx 1000\rho$) and dust particles ($\rho_p \approx 3000\rho$) but is very rarely relevant when the carrier fluid is water.

3.2.6 History term and stopping time

The Stokes time scale for plankton is only a few milliseconds at most. For an adult copepod such as *Acartia tonsa*, $\tau \approx 30 \times 10^{-3}s$, while for a flagellate such as *Bodo designis* it would be much less, $\tau \approx 25 \times 10^{-6}s$. This would thus suggest that a zooplankter, when it stops swimming, should come to a stop very quickly, if not immediately.

This is not, however the full story. An organism moving through a fluid, moves not just itself, but also a considerable amount of fluid within its boundary layer. It is the dynamics of this boundary layer, how it develops and interchanges momentum between the fluid and the moving body that is the source of the the Basset-Boussinesq history force in Eq. 3.2.24. In particular, the dynamics of the boundary layer isolate a moving body somewhat from rapid exchanges of momentum.

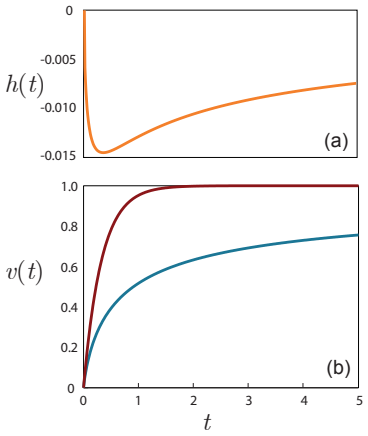


Figure 3.9: The effect of Basset-Boussinesq history force in the acceleration of a spherical particle under the influence of an applied body force of magnitude f, initiated at $t = 0$ (t in units of the Stokes time scale) on a body initially at rest. The history term is shown in (a) (in units of f), and the velocity in (b) (in units of the terminal velocity). The velocity without the history term is plotted in red, while that with the history term is in blue.

To examine this, we can can look at the case where a spherical particle, initially at rest in a motionless fluid ($u = 0$), suddenly comes under the influence of a steady body force f. The corresponding Maxey-Riley-Gatingol equation, including the history term, can be written as.

$$\alpha\frac{dv}{dt} = f - \lambda v - \lambda\int_0^t \frac{1}{\sqrt{\pi(t-t')}}\frac{dv}{dt'}dt' \tag{3.2.45}$$

The solution v' for the case where the history term is ignored is given

$$v'(t) = v_0\left(1 - e^{-t\lambda/\alpha}\right) \tag{3.2.46}$$

where $v_0 = f/\lambda$ is the terminal velocity of the particle under the action of f. The fully solution, including the history term is solved numerically, and presented in Fig. 3.9. While $v'(t)$ shows a rapid transition(on a time scale τ), the full solution $v(t)$ approaches terminal

velocity much more slowly (Fig. 3.9). With the history term included, the realized velocity has only reached 80% of terminal velocity after 5τ. While this example is posed in terms of an applied force that is turned on suddenly, the same result applies for a propulsive force that is suddenly turned off. An organisms continues to coast for a period of many Stokes time scales before it fully comes to a stop.

3.2.7 Preferential concentration of plankton by unsteady flows

One issue that surfaces now and then is the preferential concentration of plankton by small-scale, unsteady flows such as turbulence. Specifically, the concept is that coherent structures in a flow field may cause drifting particles such as plankton and detritus to diverge from the flow lines and aggregate within certain regions. At larger scales, such phenomena are fairly common, although they arise through different mechanisms. For instance, buoyant particles carried by horizontally convergent flows cause aggregations along fronts or within Langmuir circulation cells[7].

At smaller scales, preferential concentration could potentially impact fundamental processes such as trophic interactions between planktonic predators and their prey, or the rate at which detrital material coalesces and sink.

Two distinct mechanisms are identified. The first relates to the *density contrast* between planktonic particles and the fluid within which they are imbedded. This magnitude of this effect is proportional to $(1-\alpha)St$ where $\alpha = \beta^{-1}$ is a parameter expressing the density contrast and St is the Stokes' number. For neutrally buoyant particles $\alpha = 1$, and this effect disappears. The other potential mechanism relates to spatial structures in the flow field. I term this *hyperbolic ejection*. The case for this mechanism is much less secure as it only comes into effect for $St > 1$, a condition that violates the legitimacy of underlying dynamic analysis.

Density contrast: Sphere in an oscillating flow: To begin with, let us examine a case, that of an oscillating fluid flow that is spatially uniform of the form

$$u(t) = U\cos\omega t \tag{3.2.47}$$

Here U is the magnitude of the flow, and ω is the frequency. Since there are no gradients in the flow field, $D/Dt \equiv d/dt$. With gravity "turned off", Eq. (3.2.39) yields

$$\alpha\frac{dv}{dt} = \frac{du}{dt} + \lambda(u-v) \tag{3.2.48}$$

the general solution of which is

$$v(t) = U\frac{\lambda^2+\alpha\omega^2}{\lambda^2+\alpha^2\omega^2}(\cos\omega t - e^{-t\lambda/\alpha}) + U\frac{(\alpha-1)\lambda\omega}{\lambda^2+\alpha^2\omega^2}\sin\omega t + v_0 e^{-t\lambda/\alpha} \tag{3.2.49}$$

Here v_0 is the initial velocity of the particle. Note also $\alpha = \beta^{-1}$ and $\lambda = (\beta\tau)^{-1} = 3\eta/a^2$. When the particle is neutrally buoyant($\rho_p = \rho \Rightarrow \alpha = \beta = 1$), the particle motion becomes

$$v(t) = U\cos\omega t + (v_0 - U)e^{-t\lambda} \tag{3.2.50}$$

[7]Counter-rotating vertical circulation cells set-up by an interaction between wind, waves and planetary rotation, generally of the scale of 10's of metres and nearly parallel to the wind direction. Accumulations of oil and floating debris along the convergence zone between the cells leads to wind-rows on the ocean's surface.

That is, there is a decaying transient with time scale $\beta\tau = \lambda^{-1}$ that arises due to any initial velocity difference between the particle and the fluid. Once this transient has decayed, the particle motion exactly matches the motion of the fluid, a characteristic that is independent of the frequency of the fluid oscillation.

If, on the other hand, we assume *a priori* that the particles have a density far in excess of the fluid ($\rho_p \gg \rho$, $\alpha \approx (2/3)\rho_p/\rho$), and with gravity "turned off" again, Eq. (3.2.39) reduces to

$$\alpha\frac{dv}{dt} = \lambda(u - v) \tag{3.2.51}$$

With the fluid flow prescribed by Eq. 3.2.47, the motion for this heavy-particle approximation is given by:

$$v(t) = U\frac{\lambda^2}{\lambda^2 + \alpha^2\omega^2}(\cos\omega t - e^{-t\lambda/\alpha}) + U\frac{\alpha\lambda\omega}{\lambda^2 + \alpha^2\omega^2}\sin\omega t + v_0 e^{-t\lambda/\alpha} \tag{3.2.52}$$

Again we can see there is a decaying transient depending on any initial velocity differences. In the limit where these transients have decayed, the particle velocity is

$$v(t) = U\frac{1}{1 + \alpha^2 St^2}\cos\omega t + U\frac{\alpha St}{1 + \alpha^2 St^2}\sin\omega t \tag{3.2.53}$$

where $St = \omega/\lambda$. The amplitude of the velocity difference between the fluid and the particle (slip velocity)– assuming that the particle is heavy – is given by:

$$|w^*| = \frac{\alpha^2 St^2}{1 + \alpha^2 St^2} \tag{3.2.54}$$

while that for any density contrast is given by.

$$|w| = \frac{(\alpha - 1)^2 St^2}{1 + \alpha^2 St^2} \tag{3.2.55}$$

While it is clear that these become the same as $\alpha \gg 1$ (heavy particle limit), it is also clear that they are considerably different as $\alpha \to 1$ (neutrally buoyant limit). Much of the literature concerning particle motion in a fluid was built up on the behaviour of aerosols and dust particle in the atmosphere – that is for the case where the particles are considerably heavier (1 to 2 thousand times) than the carrier fluid. In its application to plankton, this fact is sometimes forgotten, and results that were originally derived for heavy particles are misapplied to settings where particle and fluid are of similar density.

Density contrast: Sphere in a rotating flow: The basic concept of preferential concentration of plankton by eddies is straight forward – heavy particles ($\rho_p > \rho$) in a rotating flow migrate outwards, while light particles ($\rho_p < \rho$) migrate inwards. This is sometimes referred to a an inertial effect, a label that is misleading. This outward or inward migration is caused by centripetal acceleration, much in the manner of a centrifuge. To orbit, a fluid element experiences a pressure gradient force directed in towards the centre of the eddy – a pressure gradient that also acts on any particles imbedded in the fluid.

Hyperbolic ejection: One subtle effect that has been proposed as a mechanism for the preferential concentration of particles, has to do with spatial structures in flow fields and the slight but important distinction between what is meant by fluid and particle accelerations: D/Dt

and d/dt. This mechanism potentially acts on all particles, even those that are neutrally buoyant, and hence of some interest for plankton ecology. In order to distinguish it from the density contrast effect discussed above, I term this *hyperbolic ejection* as it is associated with hyperbolic points in the flow field; regions where the flow exhibits strong convergence and divergence.

To start, we can pretty much discount uniform acceleration as a mechanism for the preferential concentration of neutrally buoyant particles irrespective of how subtle the mathematical formulation of the Maxey-Riley-Gatingol equations. Consider a sealed, fluid-filled container with no air pockets, in which floats a neutrally buoyant body, and the fluid being incompressible. The first thing to note is that the neutral buoyancy of the body does not depend on the strength of gravity. Doubling, or halving gravity, or for that matter allowing gravity to fluctuate over time, will have no impact on the motion of the body with respect to the fluid. It will remain neutrally buoyant at all times, suspended in the fluid with no relative motion. By the same token, any spatially uniform acceleration, no matter its magnitude or variability, will have no effect — Einstein's general relativity again (Sec. 3.2.2). Neither will the size nor the shape of the suspended body change this. Thus, any physical mechanism that leads to the divergence of the path of a neutrally buoyant particle from that of an ideal tracer or fluid element (*i.e.* crossing streamlines), must depend on the spatial structure of a time varying flow field, and not solely on the finite size and inertia of the suspended particle.

For a neutrally buoyant sphere ($\beta = 1$), Eq. 3.2.39 can be written

$$\frac{d\mathbf{v}}{dt} = \frac{d\mathbf{u}}{dt} + (\mathbf{u} - \mathbf{v}) \cdot \nabla \mathbf{u} - \lambda(\mathbf{v} - \mathbf{u}) \tag{3.2.56}$$

Writing this in terms of the velocity difference between the particle and the fluid $\mathbf{w} = \mathbf{v} - \mathbf{u}$, this becomes

$$\frac{d\mathbf{w}}{dt} = -(\mathbf{J} + \lambda \mathbf{I}) \cdot \mathbf{w} \tag{3.2.57}$$

where $\mathbf{I}$ is the identity matrix, an $\mathbf{J}$ is the jacobean of the undisturbed flow given by

$$\mathbf{J} = J_{ij} = u_{i,j} = \begin{pmatrix} \partial_x u_x & \partial_y u_x \\ \partial_x u_y & \partial_y u_y \end{pmatrix} \tag{3.2.58}$$

where I have assumed a 2-D flow field for convenience. The general form of the equation can be diagonalized to read

$$\frac{d\mathbf{w}'}{dt} = \begin{pmatrix} \gamma - \lambda & 0 \\ 0 & -\gamma - \lambda \end{pmatrix} \cdot \mathbf{w}' \tag{3.2.59}$$

Thus, if $\mathrm{Re}(\gamma) > \lambda$, an initially small mis-match between the particle velocity and the fluid flow can grow exponentially leading to a separation of trajectories. Large values of $\mathrm{Re}(\gamma)$ are characteristic of hyperbolic points in the flow: "stagnation" points if you will, where fluid is squeezed and stretched in different directions. Hence my terminology of hyperbolic ejection. An analysis of this mechanism was presented by Babiano et al. (2000) who examined the effect in a 2D steady flow field defined by the stream function

$$\psi(x, y) = A \cos(x/L) \cos(y/L) \tag{3.2.60}$$

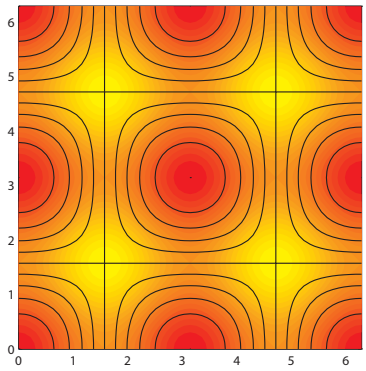


Figure 3.10: An idealized eddy field. The stream lines ψ (Eq. (3.2.60)) are indicated in black and the colours indicate values of γ^2 (Eq. (3.2.63)) – red indicates negative and yellow indicates positive values. Wherever γ is real and greater than λ neutrally buoyant particles can be propelled across streamlines. This happens in yellow area, close to hyperbolic points in the flow.

(Fig. 3.10)where the velocity field is given by

$$u_x = \partial_y \psi \tag{3.2.61}$$
$$u_y = -\partial_x \psi \tag{3.2.62}$$

This flow field is non-divergent, and the critical parameter γ is given by

$$\gamma^2 = -\det \mathbf{J} = (\sigma^2 - \omega^2)/4 \tag{3.2.63}$$

where σ is the strain and ω is the vorticity; γ is real and positive when $\sigma > \omega$ *i.e.* at hyperbolic points. The ensuing trajectory of particles released in such a flow field become chaotic (Babiano et al., 2000) even though the underlying flow structure is perfectly deterministic. Mathematically fascinating no doubt, but does it have any bearing on a natural process.

Let us go back, to the original condition, that $\mathrm{Re}(\gamma) > \lambda$. Since it is composed of velocity gradients, the scale of $\gamma \sim U/L$ where U is the velocity scale and L the length scale of the fluid flow. Thus what this condition actually states is that this effect (*i.e.* hyperbolic ejection of inertial particle) only come into effect when the Stokes' number $St > 1$. This is in direct conflict with one of the condition (Eq. (3.2.29)) under which the dynamic analysis (Eq. (3.2.24) and derivatives there of) are valid, namely that $St \ll 1$!! While Babiano et al. (2000) use a nominal Stokes' number $St = 0.2$, marginally compliant with Eq. (3.2.29), they place it in a flow field of the form Eq. (3.2.60) with $A = 100$. Since $U \propto A$, the actual Stokes' number they use is $St = 20$, well outside the validity criterion.

If hyperbolic ejection (*i.e.* the mechanism leading to preferential concentration of perfectly neutrally buoyant particles) is evident in models, then in all likelihood it is due to an inappropriate application of the underlying dynamic description, rather than a real physical effect. Only density contrast is a credible mechanism for the preferential concentration of near-neutrally buoyant particles such as plankton. There is no empirical evidence to date that neutrally buoyant particles in spatially variable flows such as turbulence exhibit preferential concentration. They remain well mixed.

3.3 Multi-pole expansion

The specific hydromechanical disturbance generated by a plankter depends on the sum of the forces it experts on the fluid (specifically thrust and drag), their spatial distribution, the boundary conditions imposed by the solid body of the organisms itself, and the Reynolds

number regime in which the flow is generated. For Stokes flow (strictly for $Re \approx 0$ but generally applicable for $Re < 1$), exact solutions for simple body shapes and force distributions can often be found. Furthermore, because the dynamics of Stokes flow are linear, solutions for more complex systems can be found by adding together solutions of simpler systems. That is solutions to complex systems can be built up by the super-position of simpler solutions.

3.3.1 Stokeslet, stresslet and rotlet

The multi-pole expansion is based on this principal of superposition. A specific multi-pole is a Stokes flow solution to a fundamental force distribution. The most fundamental of these is the stokelet – a mono-pole – and describes the flow generated by a single point force acting on the fluid. Next in complexity come the stresslet and rotlet - dipoles - the flows generated by two point forces separated by a small distance; the stresslet when the offset is parallel to the direction of forces, and rotlet when the offset is perpendicular. And so forth to quarda-poles and higher.

The central theme of this section is that complex hydromechanical flows can be considerably simplified using multi-pole expansions and to shed light on gross flow characteristics. This has several useful applications in the way we can analyze the flows generated by swimming, sinking or feeding plankton. In many instances it is the far-field characteristics of the flow that is of significance; as in the remote detection of predators and prey for instance. I include two somewhat more technical subsections; the first of which examines the flow generated by a point force in an inviscid fluid. The solution of this problem relies of the application of Green's functions (Sec. 3.3.2) where a derivation and discussion is included for the interested reader. There follows an introduction to spherical harmonics, (Sec. 3.3.3), functions that can be used to obtain flow field solutions to various problems related to the motion of spherical bodies in fluids.

Stokeslet: The solution of the Navier-Stokes' equations for an incompressible fluid at zero Reynolds number subject to a point force $\mathbf{f} = 8\pi\mu\boldsymbol{\alpha}$ is

$$\mathbf{u}_1(\mathbf{x}, \boldsymbol{\alpha}) = \frac{\boldsymbol{\alpha}}{r} + \frac{(\boldsymbol{\alpha} \cdot \mathbf{x})\mathbf{x}}{r^3} \tag{3.3.1}$$

$$p_1(\mathbf{x}, \boldsymbol{\alpha}) = 2\mu \frac{\boldsymbol{\alpha} \cdot \mathbf{x}}{r^3} \tag{3.3.2}$$

for the velocity $\mathbf{u}$ and pressure p fields (the factor $8\pi\mu$ is just a convenience).

The fundamental symmetry of the problem – that the flow and pressure fields will appear identical in any plane that is parallel with the force vector – means that spherical polar coordinates can be conveniently used. The associated stream-function in (r, θ) is:

$$\psi_1(r, \theta) = \frac{3}{4} U a r \sin^2\theta \tag{3.3.3}$$

and the velocity field can be regained by the application of

$$u_{1r} = \frac{1}{r^2 \sin\theta} \frac{\partial \psi_1}{\partial \theta} = \frac{3}{2} \frac{Ua}{r} \cos\theta \tag{3.3.4}$$

$$u_{1\theta} = \frac{1}{r \sin\theta} \frac{\partial \psi_1}{\partial r} = \frac{3}{4} \frac{Ua}{r} \sin\theta \tag{3.3.5}$$

where u_r and u_θ are the radial and zenith directions respectively.

Here I have used a convenient rescaling is to write $\mathbf{g} = (3/4)a\mathbf{U}$ where a is the size of a moving body and $\mathbf{U}$ is its velocity. Thus

$$\mathbf{f} = 8\pi\mu\boldsymbol{\alpha} = 6\pi\mu a\mathbf{U}$$

is the drag on a moving sphere. Thus $\mathbf{f}$ and $\boldsymbol{\alpha}$ can be related to plankton kinematics.

This flow field (Eq. 3.3.3) is termed a stokeslet in recognition of its fundamental nature, and the role George Stokes played in the foundations of fluid mechanics. Because the dynamics of Stokes' flows are linear, in principal any Stokes flow can be constructed by adding stokeslets together. This approach has been used in analyzing for instance, flagellated swimming (Lighthill, 1975; Childress, 1981) and the generation of feeding currents (Childress et al., 1987; Jiang et al., 1999).

For simple body shapes and force distributions, the flow field can also be modelled by a Taylor expansion of a stokeslet at the geometric centre of a body, and evaluated on its surface. This leads to the so-called multipole expansion of Stokes flows (Kim and Karrila, 1991; Leal, 1992)). In this treatment, the flow can be considered as the superposition of successively higher order 'poles' such that they reflect the appropriate dynamics and boundary conditions. The term 'poles' is borrowed from classical electro-magnetism where, for instance, the magnetic field associated with the north and south poles of a magnet is described by a magnetic (potential) dipole. In the present context, the pole hierarchy enumerates the forces acting on the fluid. That is, the monopole arises from a single force, and is equivalent to the stokeslet. Two forces gives rise to a dipole, four to a quadrapole and so forth. In general, the velocity field, u_n, associated with successively higher 'poles' scales as

$$u_n \approx U\left(\frac{a}{r}\right)^n \tag{3.3.6}$$

where U is a velocity scale (*e.g.* swimming speed, feeding current velocity, settling velocity), a is a length scale (*e.g.* size of the plankter), r is radial distance from the origin of the disturbance and n corresponds to the 2^{n-1} pole. The fact that this expansion is valid for $n \geq 1$ reflects the physical constraint that the fluid disturbance vanishes at $r \to \infty$.

The general principal here is that if we can identify the lowest order, n, associated with a particular behaviour of a plankter, we have come a long way in specifying the hydromechanical flow characteristics associated with this behaviour. This is because higher order poles, irrespective of how they are distributed, attenuate more rapidly with distance from the source.

Stresslet and Rotlet: The second order multipole is the force dipole. Physically this represents the coupled action of two forces of equal and opposite strength (*i.e.* $\mathbf{f}$ and $-\mathbf{f}$), separated by the vector $\mathbf{b}$ (Fig. 3.11b). The flow thus generated is defined by

$$\mathbf{u}_2(\mathbf{x}, \boldsymbol{\alpha}) = (\mathbf{b} \cdot \nabla)\mathbf{u}_1(\mathbf{x}, \boldsymbol{\alpha}) \tag{3.3.7}$$

where ∇ is the divergence operator and its specific form for spherical polar co-ordinates can be found in appendix 2 of Batchelor (1967). The general dipole can be best described as the sum of two components. These are the stresslet:

$$\mathbf{u}_{2\Updownarrow}(\mathbf{x}, \boldsymbol{\alpha}, \mathbf{b}) = \left(-\frac{\boldsymbol{\alpha} \cdot \mathbf{b}}{r^3} + 3\frac{(\boldsymbol{\alpha} \cdot \mathbf{x})(\mathbf{b} \cdot \mathbf{x})}{r^5}\right)\mathbf{x} \tag{3.3.8}$$

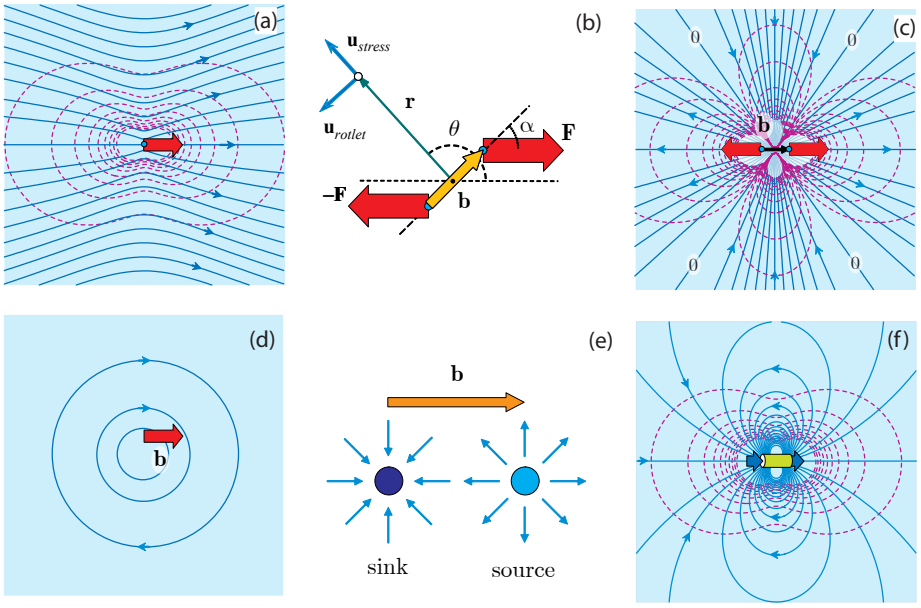


Figure 3.11: Stream functions (blue) and isotachs associated with some fundamental mulipoles; (a) the stokeslet, (c) the stresslet, (d) the rotlet and (f) the potential dipole.

and the rotlet:

$$\mathbf{u}_{2\circlearrowleft}(\mathbf{x},\boldsymbol{\alpha},\mathbf{b}) = \frac{(\mathbf{b}\times\boldsymbol{\alpha})\times\mathbf{x}}{r^3} \tag{3.3.9}$$

The flow associated with the general force dipole is the sum of these *i.e.* $\mathbf{u}_2 = \mathbf{u}_{2\updownarrow} + \mathbf{u}_{2\circlearrowleft}$. For the stresslet (Eq. 3.3.8), the flow is always radial, directed towards or away from the origin. On the other hand, the rotlet flow (Eq. 3.3.9) is always directed perpendicular to the radial direction and circulates around the origin. The nature of the flow can best be seen for the two cases where $\mathbf{f}$ and $\mathbf{b}$ parallel, and where $\mathbf{f}$ and $\mathbf{b}$ perpendicular. In the first case ($\mathbf{f}$ and $\mathbf{b}$ parallel, $\xi = 0$ in Fig. 3.11b), $\mathbf{u}_{2\circlearrowleft} = 0$, and $\mathbf{u}_{2\updownarrow}$ is a deformation flow, given in spherical polar co-ordinates (Fig. 3.5) by

$$\mathbf{u}_{2\parallel}(\mathbf{x},\boldsymbol{\alpha},\mathbf{b}) = |\boldsymbol{\alpha}||\mathbf{b}|(3\cos^2\theta - 1)\mathbf{x} \tag{3.3.10}$$

and the stream function

$$\psi_{2\parallel} = -\frac{3}{4}Ua^2\sin^2\theta \tag{3.3.11}$$

from which the radial and zenith velocity components (u_r, u_θ) can be retrieved via Eq. 3.3.4 and Eq. 3.3.5. In the second case ($\mathbf{f}$ and $\mathbf{b}$ are perpendicular), $\mathbf{u}_{2\updownarrow}$ is again a deformation flow (similar to Fig. 3.11c but rotated by 45°) and $\mathbf{u}_{2\circlearrowleft}$ is a pure rotational flow shown in Fig. 3.11d.

The potential dipole: One final multipole of relevance to this work is the potential (or source) dipole with flow given by

$$\mathbf{u}_{4\otimes}(\mathbf{x},q,\mathbf{b}) = q\left(-\frac{\mathbf{b}}{r^3} + 3\frac{\mathbf{x}(\mathbf{b}\cdot\mathbf{x})}{r^5}\right) \tag{3.3.12}$$

with associate streamfunction

$$\psi_{4\otimes} = \frac{1}{2} U \frac{a^3}{r} \sin^2\theta \tag{3.3.13}$$

Physically, this can be seen as the flow produced by a fluid source of strength $Q = 4\pi q$ (a volume flux, units m^3/s) and a fluid sink of equal strength separated by the vector $\mathbf{b}$ (Fig. 3.11.e). Indeed, we can imagine this as a short section of pipe, length $|\mathbf{b}|$, sucking water in at one end at a rate Q, and expelling it at the other. This analogy however is limited in that the inertia of fluid that is expelled is not simply in the direction of the pipe, but is radially symmetric, like an inverted siphon. The stream function associated with the potential dipole is plotted in Fig. 3.11.f. An important feature of the potential dipole is that it has zero vorticity everywhere. It is this feature that makes the potential dipole important not just for low Re, but also at high Re. While this flow is commonly termed the potential (or source) dipole, it is equivalent to a force quadrapole. This nomenclature can be confusing, and I will try to be consistent in using the appropriate qualifier (force or source or potential) when speaking of specific dipoles.

Characteristics: An important point to note is that each of the listed multipoles (stokeslet, the two force dipoles: stresslet and rotlet, and the force quadrapole or potential dipole) describes a unique lowest order property of the net force distribution or flow field. Specifically, the stokeslet exerts a single net force whereas the stresslet exerts no net force on the fluid. The rotlet on the other hand is the lowest order multipole that exerts a net torque on the fluid while the potential dipole generates a flow field with zero vorticity. Inversely, if we can identify one of these specific physical properties as characterizing a planktonic motility behaviour, we can deduce the lowest order flow field description associated with it. That is, for the 4 cases

- non-zero mean force $\Leftrightarrow$ stokeslet
- zero mean force $\Leftrightarrow$ stresslet
- non-zero mean torque $\Leftrightarrow$ rotlet
- irrotational flow field $\Leftrightarrow$ potential dipole

3.3.2 Green's functions

Green's functions are used to solve inhomogeneous differential equations subject to some specified boundary conditions. They are named after George Green (1793 – 1841) an enigmatic British mathematician who, apparently self-taught, wrote a ground-breaking treatise entitled "An essay on the application of mathematical analysis to the theory of electricity and magnetism". This treatise was written above his father's windmill on the outskirts of Nottingham. Aged 40, Green entered Cambridge University as an undergraduate, graduated in 1837 and stayed on to conduct research into optics, acoustics and hydrodynamics, but died in 1841. His works were largely ignored until discovered by William Thomson (Lord Kelvin).

In modern terminology, Green's functions are the solution of a linear operator $\mathfrak{L}$ subject to an singular impulse, mathematically written in terms of the Dirac delta function δ. That is

$$\mathfrak{L}\{G(x,s)\} = \delta(x-s) \tag{3.3.14}$$

where s is the location of the impulse, x is the observation location, and $G(x,s)$ is the Greens function associated with $\mathfrak{L}$. Following on from this, if we have a problem of the form $\mathfrak{L}\{u(x)\} = f(x)$ then its solution $u(x)$ is given by

$$u(x) = \int G(x,s) f(s) ds \tag{3.3.15}$$

Physically what this means, is that because $\mathfrak{L}$ is linear, any solution can be built-up by the superposition of elemental solutions $G(x,s)$ distributed in space in accordance with the impulse distribution $f(x)$. This fundamental concept has found broad application in whole range of scientific disciplines including electro-magnetism, quantum mechanics and hydrodynamics.[8]

One particularly important Greens function is that associated with the Laplacian operator, *i.e.* $\mathfrak{L} = \nabla^2$, that is $\nabla^2\{G(\mathbf{x},\mathbf{s})\} = \delta(\mathbf{x}-\mathbf{s})$. In 3 dimensional space, this is given by

$$G(\mathbf{x},\mathbf{s}) = -\frac{1}{4\pi r} \tag{3.3.16}$$

where $r = |\mathbf{x}-\mathbf{s}|$ is the distance from the singularity. Rearranging these results, also implies that

$$\delta(\mathbf{x}-\mathbf{s}) = -\frac{1}{4\pi}\nabla^2\left(\frac{1}{r}\right) \tag{3.3.17}$$

a useful identity that I will rely on later[9].

The divergence theorem: The volume integral of the divergence of a vector field is equal to the vector flux entering the volume through its surface.

$$\int_V \nabla\cdot\mathbf{v}dV = \int_S \mathbf{v}\cdot d\mathbf{s} \tag{3.3.18}$$

Intuitively, this can be understood in terms of an incompressible fluid; the divergence volume integral being a sum of all the sources and sinks within the volume, which must be balanced by the flow of fluid across the boundary of the volume.

Stokeslet as a Greens function: Consider the Stokes equation forced by a point force $\mathbf{f}$ acting at a location in space $\mathbf{x}_0$. We seek spatially dependent solutions p and $\mathbf{u}$ such that

$$\nabla p - \mu\nabla^2\mathbf{u} = \mathbf{f}\delta(\mathbf{x}-\mathbf{x}_0) \tag{3.3.19}$$

In particular, we seek solutions of the form

$$u_i(\mathbf{x}) = \frac{1}{8\pi\mu}\mathcal{G}_{ij}(\mathbf{x})f_j \tag{3.3.20}$$

For an incompressible fluid $\nabla\cdot\mathbf{u} = 0$, the Navier-Stokes' equation implies the pressure field is $\nabla^2 p = \mathbf{f}\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_0)$, from which using Eq. (3.3.17), it immediately follows that

$$p = -\frac{1}{4\pi}\mathbf{f}\cdot\nabla\left(\frac{1}{r}\right) \tag{3.3.21}$$

Substituting back into Eq. (3.3.19), we thus have

$$\nabla^2\mathbf{u} = -\frac{1}{4\pi\mu}\mathbf{f}\cdot(\nabla\nabla-\nabla^2)\left(\frac{1}{r}\right) \tag{3.3.22}$$

[8] A particularly useful application is for diffusion. Armed with the solution of the diffusive spread of material from an instantaneous point source, solutions to more complex configurations *e.g.* releases of material over complex-shaped source regions and/or with time-varying release rates, can be built up by spatially and/or temporal integration of the point source solution.

[9] If you are unconvinced by this line of argument, you can try it the other way round. You can readily show that $\delta(\mathbf{x}-\mathbf{s}) = \nabla^2(4\pi r)^{-1}$ so that Eq. 3.3.17 is true, from which Eq. 3.3.16 follows.

Writing $\mathbf{u} = \mu^{-1}\mathbf{f} \cdot (\nabla\nabla - \nabla^2)\mathcal{H}$ we thus have

$$\mathbf{f} \cdot (\nabla\nabla - \nabla^2)\left(\nabla^2\mathcal{H} + \frac{1}{4\pi r}\right) = 0 \tag{3.3.23}$$

This is satisfied by the Poisson equation $\nabla^2 H = -1/(4\pi r)$. Further, it can be noted via Eq. 3.3.17 that the $\nabla^4 H = -1/(4\pi r)$, the solution of which is $\mathcal{H} = r/(8\pi)$ so that

$$u_i = \frac{1}{8\pi\mu}\mathcal{G}_{ij}(\mathbf{x})f_j \tag{3.3.24}$$

where

$$G_{ij} = \frac{\delta_{ij}}{r} + \frac{x_i x_j}{r^3} \tag{3.3.25}$$

This is the free-space Green's function of the Navier Stokes' equation for a singular force. This is also known as the Orseen-Burger tensor, or simply the stokeslet.

3.3.3 Spherical harmonics and multipoles

In the Stokes regime, the equations of motion for a fluid are linear. For a large class of problems for a fluid interacting with a solid, rigid particle, this is simplified even further by noting that some physical settings have an axial symmetry; for instance flows interacting with a rigid spherical particle. Under these conditions, the flow can be characterized by a stream function ψ, which describes the trajectory of fluid elements in space. The velocity field in spherical polar coordinates then is given by

$$u_r = \frac{1}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta} \tag{3.3.26}$$

$$u_\theta = -\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r} \tag{3.3.27}$$

The momentum equations can then be expressed as

$$\mathcal{E}^2\left(\mathcal{E}^2\psi\right) = 0 \tag{3.3.28}$$

where $\mathcal{E}$ is the operator

$$\mathcal{E}^2 = \frac{\partial^2}{\partial r^2} + \frac{\sin\theta}{r^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\right) \tag{3.3.29}$$

Solutions to particular systems can then be found by solving Eq. 3.3.28 subject to specific boundary conditions, *i.e.* how $\mathbf{u}$ behaves for $r \to \infty$ and $r = a$. General solutions to Eq. 3.3.28 can be found in terms of solid spherical harmonics (Lamb, 1932) – functions of r and θ that have specific properties. A full development along these lines is not appropriate here, instead some specific solutions of immediate interest are:

$$\phi_{n,0} = \left(\frac{a}{r}\right)^n \sin^2\theta \qquad n = -2, -1, 1 \tag{3.3.30}$$

$$\phi_{n,1} = \left(\frac{a}{r}\right)^n \cos\theta\sin^2\theta \qquad n = 0, 2 \tag{3.3.31}$$

To show that these are solutions analytically can take pages of algebra. Symbolic logic programs like Mathematica, can, however, be used to quickly verify that these satisfy Eq.

3.3.28. Because the Stokes' regime is linear, any superposition of these solutions is also a solution. Thus mixing and matching these solutions so that they comply both with the physical force balance and with the boundary conditions imposed by a particular situation will give a particular full solution.

A sinking sphere: For a sinking particle, we know that the net force balance on the fluid is a stokeslet, corresponding to $\phi_{-1,0}$ above. Further, we know that the particle moves, so that the far field flow relative to the particle also involves a $\phi_{-2,0}$ contribution. Specifically

$$\psi \approx -\frac{U}{2} r^2 \sin^2\theta \tag{3.3.32}$$

for $r \to \infty$. Finally, in order for the boundary condition of zero flow at the surface of the particle to be satisfied, a $\phi_{1,0}$ contribution is also required. Thus the general solution should look like

$$\psi = -\frac{U}{2}\left(r^2 + c_1 a r + c_2 \frac{a^3}{r}\right)\sin^2\theta \tag{3.3.33}$$

The boundary condition that $u_r = 0$ at $r = a$ means that $1 + c_1 + c_2 = 0$, and $u_\theta = 0$ at $r = a$ requires that $2 + c_1 - c_2 = 0$. Thus, $c_1 = -3/2$ and $c_2 = 1/2$, and the full solution is

$$\psi = -\frac{U}{2}\left(r^2 - \frac{3}{2} a r + \frac{1}{2}\frac{a^3}{r}\right)\sin^2\theta \tag{3.3.34}$$

This corresponds exactly with the low Reynolds' number flow past a sphere as derived in Sec. 3.2.1 by a more circuitous route. This technique will be used later in deriving the family of solutions for the flow around a self propelled sphere (Sec. 3.4.4).

3.4 Hydrodynamics of plankton propulsion

Because of their small size and relatively slow speeds, plankton generally experience a low Reynolds number regime. That is the frictional (*i.e.* viscous) forces they experience are generally much larger than their inertia. With respect to propulsion, this means that plankton have to "crawl" through the fluid rather than rely on inertial thrust as we do, or for that matter a fish does when swimming. To best utilize the hydrodynamic regime they inhabit, plankton have developed a number of structures and abilities including flagella, cilia, paddles and squirming to provide propulsion. But how exactly do these work?

3.4.1 Mechanisms

The nature of low Reynolds number propulsion is fundamentally different from that which we experience. This is because friction dominates; any forward motion achieved during a power stroke would be exactly cancelled by backward thrust during the recovery stroke. The kick of a human swimmer, or the paddling of ducks simply would not work at low Reynolds number. The trick to low Reynolds number propulsion is to find asymmetries in the drag associated with different phases of the stroke.

As an example, a two "paddled" organism, such as that depicted in Fig. 3.12 would not gain any propulsion if it beats both paddles up and down synchronously; the upward force on the down beat would be exactly cancelled out by a downward force on the up beat. The relative speed of the beats won't matter. True, the drag is less for a slower beat, but it would have to

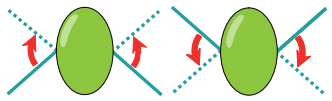


Figure 3.12: A two paddle synchronous swimmer does not have any net movement at low Re. Any motion caused by the power stroke is exactly cancelled by motion caused by the return stroke.

be sustained longer to return to the same position. The integrated force over a whole cycle will always be zero. This principle of low Reynolds' number flow is sometimes called Purcell's scallop theorem[10].

However, the same organism can move if it breaks the symmetry of its motion: right paddle down while holding its left paddle fixed, then left paddle down while holding its right paddle fixed and so forth (Fig. 3.13). This is essentially the locomotion mode of a "three-hinged" swimmer as proposed by Purcell (1977) in his seminal work on life at low Reynolds numbers. While the three-hinged swimmer, purportedly one of the simplest means of low Reynolds number locomotion, has been demonstrated to work in scale models (Hosoi and Chan), a succinct analysis of the underlying physics is elusive. Purcell himself was annoyingly tight-lipped on the subject, leaving it "as an exercise for students". Recently, Becker et al. (2003) provided a rigourous analysis, showing that while propulsion is possible, it is very inefficient, and the direction of net motion can be either left or right depending on the relative length of the rods and the amplitude of their swing.

Purcell's conjecture, despite its elusive solution, has stimulated considerable interest in recent years. This interest has arisen chiefly in the area of micro-technology and the design of micro-machines to move or move through fluids. This has lead to proposals of even simpler swimmers, ones where the underlying physics is more readily understood (*cf.* Fig. 3.14 and Fig. 3.15.

The basic problem of swimming can be stated as seeking a solution to the Navier-Stokes' equations subject to the specific boundary conditions of no slip on the time varying body shape of the swimmer. If the body shape relative to the centre of mass of the swimmer is prescribed by the function $\mathbf{u}_s(t)$ (known as the swimming gait), then the inner boundary conditions are prescribed by

$$\mathbf{u}(\mathbf{s}) = \frac{d\mathbf{s}}{dt} = \mathbf{U} + \mathbf{\Omega} \times \mathbf{s} + \mathbf{u}_s(t) \tag{3.4.1}$$

where $\mathbf{x} = \mathbf{s}(t)$ is the surface of the swimmer and $\mathbf{U}(t)$ and $\mathbf{\Omega}(t)$ are the resultant translation velocity and rotation of the swimmer as a whole. Formally, theses can be solved using the reciprocal theorem. This remarkable theorem, due to H.A. Lorentz [11], relates the resulting

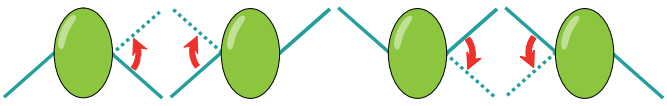


Figure 3.13: A two paddle asynchronous swimmer can move at low Re.

[10]Edward Mills Purcell (1912 – 1997) shared the Nobel prize in physics in 1952 for his work on nuclear magnetic resonance. His interest in biophysics was stimulated by his association with Howard Berg.

[11]Hendrik Antoon Lorentz (1853 – 1928), a dutch physicist who shared the 1902 Nobel prize for physics

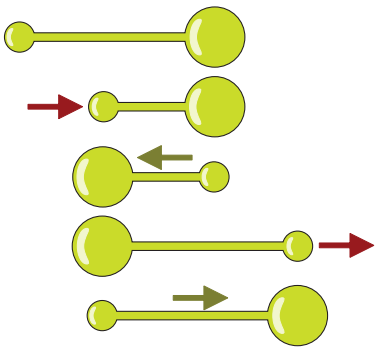


Figure 3.14: The pushmepullyou (Avron et al 2005), is a relatively simple swimming device of two inflatable spheres connected by an extensible tube through which fluid is pumped. The drag on an inflated sphere, being larger, is greater than that on a deflated sphere. By pumping up alternate spheres and extending and contracting the connecting tube, the device can translate relative to the fluid.

flows of *any* 2 different sets of applied actions on a given body shape to each other. That is, if $\hat{\mathbf{F}}$ and $\hat{\mathbf{T}}$ are an arbitrary external force and an arbitrary torque applied to a particular realization of the swimming gait $\mathbf{u}_s$ leading to the solution $\hat{\mathbf{u}}$ and $\hat{\boldsymbol{\sigma}}$, then the reciprocal theorem states:

$$\hat{\mathbf{F}} \cdot \mathbf{U} + \hat{\mathbf{T}} \cdot \boldsymbol{\Omega} = \int_{s(t)} \mathbf{u} \cdot \hat{\boldsymbol{\sigma}} \cdot \mathbf{n} dS \tag{3.4.2}$$

Since $\hat{\mathbf{F}}$ and $\hat{\mathbf{T}}$ are arbitrary, there is more than enough information available to solve for all components of $\mathbf{U}(t)$ and $\boldsymbol{\Omega}(t)$. The way to understand this physically is to think of the arbitrary $\hat{\mathbf{F}}$ and $\hat{\mathbf{T}}$ as test functions whose effect on the body $\mathbf{u}_s$ is to produce some response $\hat{\mathbf{u}}$ and $\hat{\boldsymbol{\sigma}}$. But because the Navier Stokes' equations are linear, this response carries information about the effect of $\mathbf{u}_s$ for *any* force and torque.

The 2 basic issues related to Stokes' swimming are rate independence and time reversibility (Purcell's scallop theorem), both of which arise again from the linear nature of the Navier-Stokes' equations. Rate independence arises because the Navier Stokes equations have no specific time dependence. Each response to a specific action is essentially instantaneous. Like snapshots laid in sequence as a movie, the final outcome does not depend on whether the movie is played at high speed, normal speed or slow motion, or variable speed for that matter. Thus the timing of stroke whether rapid or slow makes has no influence on the net motion of a swimmer. The other aspect arising out of this is the time reversible criterion. Simply put, if the movie of actions look the same when played backwards as they do forwards, then the swimmer executing them would not experience net movement.

In nature, plankton[12] have developed a whole range of mechanical devices – whips, hairs, wings, pumps and paddles – to achieve propulsion. A collection of fascinating movies can be found at several sites: for bacteria and protist.

The 1950s saw several landmark developments in the analysis of the propulsion of micro-zooplankton. This burst of activity was initiated by G.I. Taylor (Taylor, 1951, 1952a,b) a central figure in many topics touched on in this thesis – as well as Sir James Gray, Sir James Lighthill[13] and their students and collaborators (*e.g.* Hancock, 1953; Gray and Hancock, 1955; Lighthill, 1969; Blake, 1971; Chwang and Wu, 1971; Lighthill, 1975, 1976; Higdon,

for a theoretical explanation of the Zeeman effect. He also formulated the transformations used by Einstein in his theory of special relativity. The reciprocal theorem essentially states that if $(\mathbf{u}, \boldsymbol{\sigma})$ and $(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}})$ are 2 Stokes solutions for a particular configuration, then $\nabla \cdot (\mathbf{u} \cdot \hat{\boldsymbol{\sigma}} - \hat{\mathbf{u}} \cdot \boldsymbol{\sigma}) = 0$ (Lorentz, 1896).

[12] Propulsion is not the exclusive domain of zooplankton (*i.e.* heterotrophs). Many autotrophs such as ciliates and diatoms have propulsive capabilities.

[13] Sir Michael James Lighthill (1924 –1998), British applied mathematician who worked in aerodynamics,

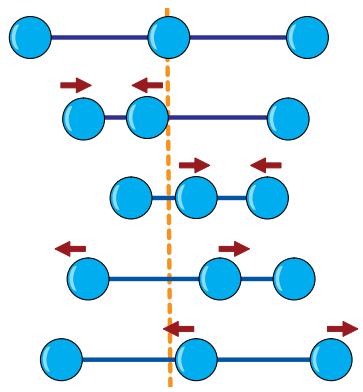


Figure 3.15: The three-sphere-creeper (Najafi and Golestnian 2004) and the three-sphere-rotator (Dreyfus et al 2005) rely on the principal that the drag felt on two spheres fixed close together is less that that when they are further apart. An animation of a race between a pushmepullyou and a three-sphere-creeper give a clear illustration of the processes involved.

1979; Lighthill, 1996). The summary papers by Lighthill are particularly well worth reading, giving a clear and detailed account of not only the physics but also the biology of micro-organisms and their locomotion. An excellent modern review is given by Lauga & Powers.

In nearly all cases, propulsion has 2 aspects; there is the movement through space, a facet we can most easily relate to, but there is also the flow of fluid close to the organism's body that propulsion effects. The latter of course facilitates the organism's proximate contact with solutes, suspended material and other biota. Swimming and feeding currents are often two aspects of the same propulsive mechanism.

3.4.2 Flagellar propulsion

Many simple organisms including protozoans, bacteria, algae and fungi have flagella – one or more whip-like organelles – that provide a means of locomotion. Traction is generally achieved by undulations propagating along the length of the flagellum and can take the form of either planar waves (a wiggling motion) or helical waves (like a turning cork screw). The underlying physics that makes flagellar propulsion possible at low Reynolds number is that the resistance to movement of a thin cylindrical body (a small section of the flagellum for instance) is nearly twice as much when it moves with its long axis perpendicular to the direction of motion compared to when it moves with its long axis parallel.

A concise account of flagellar motion as predicted by resistive theory was provided by Gray and Hancock (1955). This is illustrated in Fig. 3.16, a small section of the flagellum of length ds (this can be envisaged as a short cylindrical rod) inclined at an angle θ to the axis of propulsion, and moving with a velocity v at an angle ϕ to this axis, will experience force normal to the rod, $dN = C_N v \cos(\theta + \phi) ds$, and a force along the line of the rod, $dL = C_L v \sin(\theta + \phi) ds$. Here C_N is the drag coefficient for flow perpendicular to the rod, while C_L is that for flow parallel to it. These may well be quite different. The net force in the direction of propulsion is thus

$$\begin{aligned} dF &= dN \sin\theta - dL \cos\theta \\ &= \left((C_N - C_L) \cos\phi \sin\theta \cos\theta - (C_N \sin^2\theta + C_L \cos^2\theta) \sin\phi \right) v ds \\ &= \left(\frac{(C_N - C_L) \cos\phi \tan\theta - (C_L + C_N \tan^2\theta) \sin\phi}{1 + \tan^2\theta} \right) v ds \end{aligned} \tag{3.4.3}$$

acoustics and biophysical fluid dynamics. From 1969 to 1979, he held the chair of Lucasian professor of mathematics at Cambridge, a chair previously held by Isaac Newton (1669–1702), George Gabriel Stokes (1849–1903) and up until recently, by Stephen Hawking (1979-2009).

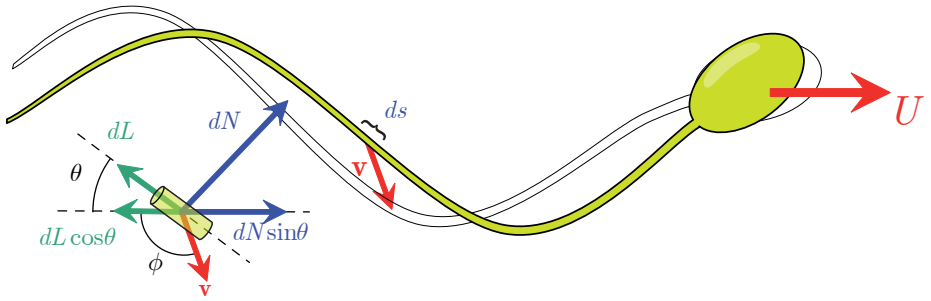


Figure 3.16: Flagellar propulsion. The tangential dN and along flagellum dL forces acting on a small section of the flagellum ds moving with respect to the fluid with velocity $\mathbf{v}$.

Now $v \sin\phi = U$, the forward motion of the organism as a whole, while $v \cos\phi = dy/dt$, the transverse speed of the oscillation relative to the moving organism. Noting that $tan\theta = dy/dx$, and assuming a sinusoidal wave form $y(x,t) = b \sin(k(x+ct))$ where the amplitude of the flagellar oscillation b is small compared to its wavelength $\lambda = 2\pi/k$, leads to

$$F = n\pi k b^2 c(C_N - C_L) - 2n\pi C_L U/k \tag{3.4.4}$$

where n is the number of waves along the flagellum. With the head (radius a, and drag coefficient C_H), the full force balance at steady state ($F = aC_H U$) gives a forward swimming speed

$$U = \frac{n\pi k^2 b^2 c(C_N - C_L)}{2\pi n C_L + kaC_H} \tag{3.4.5}$$

For a long slender cylinder (length 2ℓ, radius δ), the coefficients of friction are (approximately)

$$C_N = \frac{4\pi\mu}{\ln(4\ell/\delta) - 1/2} \tag{3.4.6}$$

$$C_L = \frac{2\pi\mu}{\ln(4\ell/\delta) + 1/2} \tag{3.4.7}$$

Using a plane wave flagella motion as a means of propulsion is not particularly efficient. This is because the direction of thrust at some points along the flagellum will be counter to that provided by the flagellum as a whole. This can be made significantly more efficient by using a helical wave (*i.e.* like a cork screw) as practiced by bacteria rotating their flagellar bundle. With this configuration the angle of attack of the thrust is always parallel to the direction of movement, so that there is comparatively little wasted effort. However, this type of propulsion does introduce a torque (*i.e.* a turning force) that will cause the organisms to rotate, and subsequently reduce the overall efficiency of forward propulsion.

3.4.3 Ciliate propulsion

Ciliates derive their name from cilia, small hair like structures that cover their bodies. Cilia are identical in structure to eukaryotic flagella in that they are formed by a bundle of microtubules, two in a central core and nine distributed radially around them. Propulsion is provided by the bending of individual cilia in a synchronized fashion so that the whole cilia

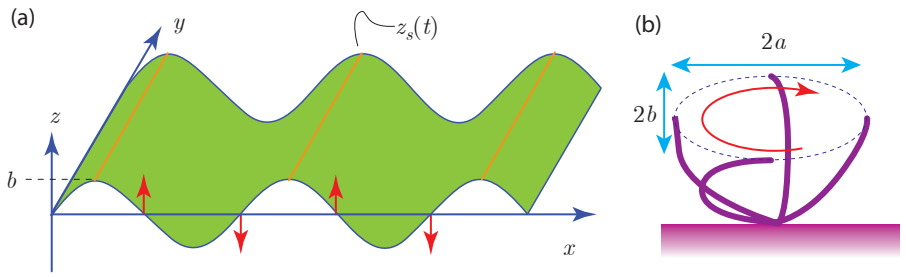


Figure 3.17: The geometry of a swimming sheet (a) seen as a deformable surface stretching infinitely in all directions on which a planar wave propagates. Wave crests travel in the $+ve$ x direction. The swimming sheet can be seen as an special case of a ciliated sheet, where the tip of each cilia traces out an ellipse whose shape depends on the amplitudes in the x and z directions (a and b respectively), and the relative phase, ϕ.

field undulates as a travelling wave across the body of the ciliate. Several treatments of the physics of ciliated motion, both the propulsive mechanism itself and its effect on the adjacent fluid, have been promoted over the years. Here I briefly discuss some of these.

The Swimming Sheet: A set of treatments of ciliated swimming begin with the concept that the oscillating cilia field can be treated essentially as a solid boundary to the fluid. One of the simplest models of swimming in this fashion is that of a 2 dimension surface immersed in a fluid, that undulates with a rippling wave. This is the swimming sheet model first proposed by GI Taylor. Consider for instance an infinite surface that ripples according to

$$z_s(t) = b\sin(kx - \varpi t) \tag{3.4.8}$$

where b is the amplitude of the ripple, k is the wave-number, and ϖ is the frequency. This represents a wave that travels in the +ve x direction. The boundary condition for the Stokes equations – assuming no slip at the solid boundary – is thus

$$\mathbf{u}(x, z_s(x,t)) = -b\sin(kx - \varpi t)\hat{\mathbf{z}} \tag{3.4.9}$$

The basic geometry is 2 dimensional in the $x-z$ plane being isotropic in the y direction Fig. 3.17. This allows a stream function ψ to be defined such that

$$u = \frac{\partial\psi}{\partial z}, \qquad w = -\frac{\partial\psi}{\partial x}$$

The Stokes' equations for an incompressible fluid in terms of the stream function reduces to $\nabla^4\psi = 0$. The boundary condition on the sheet (Eq. 3.4.9) as well as that at $z \to \pm\infty$ that $\mathbf{u} \to U\hat{\mathbf{x}}$ leads to the solution

$$U = -\frac{1}{2}\varpi k b^2 \tag{3.4.10}$$

That is, the sheet moves in the direction opposite to the direction in which the wave propagates. The solution is got by means of a small parameter expansion in kb, that the ratio of the wave amplitude to the wavelength is small(*cf.* Childress (1981) for details).

In a somewhat more realistic treatment, the tip of each cilium moves in an orbit $[x_s, z_s]$ given by:

$$x_s(t) = x + a\cos(kx - \varpi t + \phi) \tag{3.4.11}$$
$$z_s(t) = b\sin(kx - \varpi t) \tag{3.4.12}$$

where a and b are the wave amplitudes in the horizontal and vertical directions respectively, and ϕ is a phase difference between horizontal and vertical cilium tip displacements. The cilia tips are treated as an impenetrable surface with a non-slip condition. That is, the fluid on the surface $[x, z]$ moves with the exact same velocity as the cilia tips.

For an infinite 2-D oscillating sheet with displacements of the form Eq. 3.4.11, the propulsion speed is given by

$$U = -\frac{1}{2}\varpi k \left(b^2 + 2ab\cos\phi - a^2\right) \tag{3.4.13}$$

Of course swimming sheets, ciliated or not, illustrate a propulsive mechanism, but cannot be taken as representing real organisms. A more realistic representation is through an envelop model. This class of models utilize swimming sheet solutions, but confines them to the geometry of a finite 3 dimension body. That is, the skin of the organism appears to ripple with wave crests propagating rear-ward while the organism is propelled forward.

Sublayer model: The oscillating sheet and the envelope model of ciliated propulsion assume the propulsive forces arise solely due to the kinematic movement of the surface of the ciliar layer. This is somewhat simplistic as the "surface" of the ciliate layer (the tips of all the synchronously moving hairs) is far from impermeable – fluid is suck in and expelled from this layer continuously. This shortcoming has been addressed, at least in part, by the sublayer model introduced by Blake (1971). The sublayer model considers the resistive force of each cilium, in particular the resistive aspects of a cilia moving close to a solid boundary the body of the organism. Specifically, by bringing a cilium close to its body on a recovery stroke, the organism can reduce the drag compared to the fully extended cilium during the power stroke. This asymmetry in resistance propels the organism forward. Detailed modelling the of the within the sublayer can be achieved by placing a sequence of Stokeslet singularities along each cilium, an approach which is relatively accurate provided the cilia are sufficiently widely spaced (*e.g.* Blake, 1972; Lighthill, 1976).

3.4.4 The squirmer: a self propelled sphere

In Sec. 3.3.3 on page 51, I briefly discussed the formalism of spherical harmonics and how these can be used to address the problem of creeping flow. The same principle can be applied to a self propelled sphere; an idealized model for flagellate or ciliate swimming where the solid body of the organisms is represented.

By whatever means thrust is applied to the fluid, we know that the net force balance on a self propelled body is a stresslet, so that the solution should involve a function of the form $\phi_{0,1}$. But we know also that the radial component of this flow must disappear on the surface of the sphere. Thus an element of the solution is

$$\psi \approx -Ua^2\left(1 - \frac{a^2}{r^2}\right)\cos\theta\sin^2\theta \tag{3.4.14}$$

We also know that the sphere moves, so that like the sinking sphere in Sec. 3.3.3, the far field flow must approach $\psi \to Ur^2\sin^2\theta$ as $r \to \infty$. This flow must also disappear on the surface of the sphere. To include this, without dominating the known far-field stresslet characteristics, a $\phi_{1,0}$ contribution can be added. The full solution is thus

$$\psi = -\frac{U}{2}\left(r^2 - \frac{a^3}{r}\right)\sin^2\theta + \kappa\frac{3}{2}Ua^2\left(1 - \frac{a^2}{r^2}\right)\cos\theta\sin^2\theta \tag{3.4.15}$$

where κ is the so called squirmer number (Magar et al., 2003), representing the distribution of thrust and drag over the surface of the body.

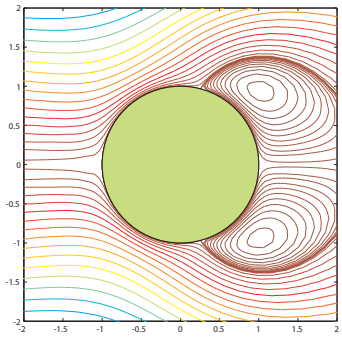


Figure 3.18: The stream function ψ for a self propelled sphere, with squirmer number $\kappa = 2$. The mean flow of fluid relative to the body in this case os to the left, or if you will, the body moves to the right with respect to the far field. This is an example of a "pusher". The separation points of the re-circulation zone forward of the body is given by $\cos\theta = 1/(2\kappa)$.

The velocity field associated with this self-propelled sphere is thus

$$u_r = U\left(\frac{a^3}{r^3} - 1\right)\cos\theta - \kappa U\frac{3}{2}\left(\frac{a^4}{r^4} - \frac{a^2}{r^2}\right)\left(3\cos^2\theta - 1\right) \tag{3.4.16}$$

$$u_\theta = U\left(1 + \frac{a^3}{2r^3}\right)\sin\theta - \kappa U\frac{3a^4}{2r^4}\cos\theta\sin\theta \tag{3.4.17}$$

which turns out to be the lowest order spherical envelope solution (Blake, 1971) that contains the stresslet force balance.

The squirmer number indicates how thrust and drag are distributed. When $\kappa = 0$, drag and thrust are distritbuted equally over the surface of the sphere so that they balance exactly at all locations. This of course is a rather special case, that would involve a refined balancing act by an organism that tried to achieve it. In general, thrust and drag are distributed asymmetrically: flagellates and bacteria are clear examples, but ciliates also likely have asymmetries in their force distributions. These asymmetries fall into 2 categories, pushers ($\kappa > 0$) where thrust is applied aft of the body (as illustrated in Fig. 3.18), and pullers ($\kappa < 0$) where thrust is applied forward of the body. While the flow field associated with a puller is the mirror image of that for a pusher, there are interestingly different effects that arise when groups of pushers and pullers swim together.

We can now calculate the power required to propel a squirmer. For a sinking particle, the power expended can be readily calculated from Stokes drag, it is the net drag times the velocity so that

$$P_o = 6\pi\mu a U^2 \tag{3.4.18}$$

For a squirmer we cannot approach this calculation in the same way as the net force applied to the fluid is zero, thrust exactly balances drag. However, we not that in any region of space, the energy dissipation rate is given by

$$\varepsilon = 2\mu e_{ij}e_{ij} \tag{3.4.19}$$

which for an axially symmetric flow field as is the case for both a sinking particle and a squirmer is given by

$$\varepsilon = 2\mu\left(e_{rr}^2 + e_{\theta\theta}^2 + e_{\phi\phi}^2 + 2e_{r\theta}^2\right) \tag{3.4.20}$$

Note that even though the azimuthal velocity component is zero, as are all gradients in that direction, the rate of strain component $e_{\phi\phi}$ is in general non-zero. Using the Stokes flow velocity components in this expression leads directly to the same result expressed in Eq. 3.4.18. For the squirmer solution, the energy dissipation and therefore the power required for propulsion becomes

$$P_s = 2\pi \int_0^{\pi} \int_a^{\infty} r^2 \sin\theta \varepsilon(r,\theta) dr d\theta = 12\pi\mu a U^2 \left(1 + 2\kappa^2\right) \tag{3.4.21}$$

That is, the power required to propel a squirmer through the water is at least twice that expended by a sinking particle of the same size moving at the same speed. This power requirement increases as the squirmer number deviates from zero.

3.4.5 Paddle propulsion and feeding currents

Copepods, some of the most ubiquitous zooplankton genera, derive their name from the Greek, *kope – podos* meaning paddle-footed. They are equipped with a row of paired swimming appendages (pereiopods). Somewhat surprisingly, most of the time, swimming appears to be facilitated by movements of the mouth parts (more properly the cephalic appendages, the maxillipeds) the use of the swimming legs being reserved for rapid bursts of speed as in an escape flight.

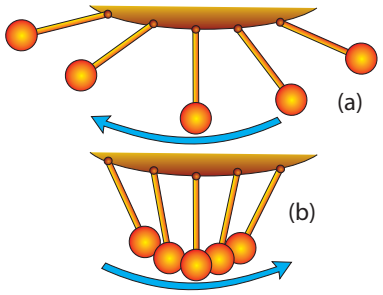


Figure 3.19: The sweep sequence of a copepods swimming legs during an escape jump: (a) during the power stroke legs kick one after the other, while (b) during the the recover stroke, legs move together.

The prominent use of maxillipeds in swimming is likely connected with a more important function of these appendages namely forming a feeding current. Copepods are generally negatively buoyant and would sink if they did not swim. When their thrust is exactly that of displaced weight, they remain motionless; gravitationally tethered if you will.

At first it may seem curious that paddling copepods can achieve any movement with respect to the fluid. A simple back and forth motion at low Reynolds number after all provides no net thrust. There are two factors however of relevance. The first is that at the size and speeds of copepod propulsion, the Reynolds number can approach and exceed unity so that inertial effects can become important. The second is that by tuning the timing and direction of these strokes relative to each other, symmetry can be broken and propulsion can be achieved. van Duren (2000) for instance showed that in a rapid escape jump, copepods kick with their swimming legs one after the other, but bring them back together as a bundle. Now recall the hydrodynamic property that allows the mechanical 3 sphere creeper to move: that the resistance on two bodies is greater when they are far apart that when they are close together. The same principle is used by copepods to break the symmetry between their power stroke and recovery stroke (Fig. 3.19).

There are other means by which this symmetry can be broken. A single appendage swept in a “figure 8” configuration (Fig. 3.20) will in general impart a net force on the fluid (Chil-

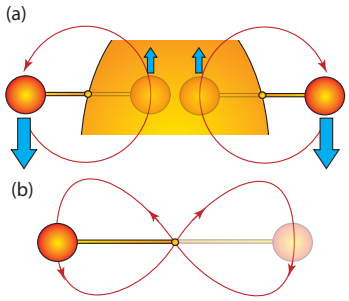


Figure 3.20: The sweep sequence of a copepods swimming legs during an escape jump: (a) during the power stroke legs kick one after the other, while (b) during the the recover stroke, legs move together.

dress et al., 1987) – recall Purcell's scallop theorem regarding time reversibility – that if the sequence of actions looks different when played backwards of forwards in time, symmetry is broken and net traction can be applied by the swimmer on the fluid. A simpler symmetry breaking motion is to sweep 2 appendages in counter-rotating circles. Such a swimming motion not only breaks symmetry, but also enjoys the reduced drag of proximate bodies as an added physical effect.

While such general principals arise out of theoretical considerations, recent high-speed microscopic recording of copepod swimming behavior have revealed that these organisms are extremely dexterous in manipulating both their own movement and the flow of fluid around their bodied. They are equipped with many paddles – not only their 10 swimming legs, but also numerous mouth parts, a tail and antennas. All of these are act in high speed manoeuvres such as catching a swimming prey. The tail flicks and rotates, the anetennae sweep down, the legs kick and then anchor while the mouth parts fling open to suck the prey in. While capture of prey at low Reynolds numbers may seem like a tricky problem to us, copepods and other zooplankton appear to be completely at home in this environment.

3.5 Hydrodynamic signals in the plankton

A plankter moving through the water, either by swimming or sinking, generates a disturbance in the surrounding fluid. In the first place, simply moving through the water causes the fluid to deform to make space for the translating organism. Secondly, at the small scales that plankton inhabit, water is viscous, and a certain amount of fluid is dragged along as it moves. Thirdly, a swimming organism must push on the fluid to get traction and move, an action that also generates a disturbance. In many instances, the fluid disturbance can be detected by other organisms (so-called rheotactic ability), either predators, prey or mates, and alert them to the proximity of a potential prey, predator or mate. That is, any motion of an organism relative to the water can generate a signal that can increase the distance at which it is detected, thus radically altering its encounter rate with rheotactic predators, prey or mates.

A great many planktonic organisms (ciliates, flagellates, copepods) exhibit rheotactic ability to a greater or lesser extent. Copepods for instance, are equipped with arrays of sensitive hairs (setae). In adult copepods (Fig. 3.21), these are arranged largely along their first antennae (Fig. 3.22) and in some species, may also be prominent on their telson (tails). Early life stages of copepods (nauplii and copepodites) also have setae more or less arranged uniformly over their body surface. These setae act as tiny shear probes. When the tip of a seta is moved by a fluid disturbance, micro-tubules in a hinge at the base of the seta are stretched, triggering a neurological response. This neurological response has been observed by inserting micro-

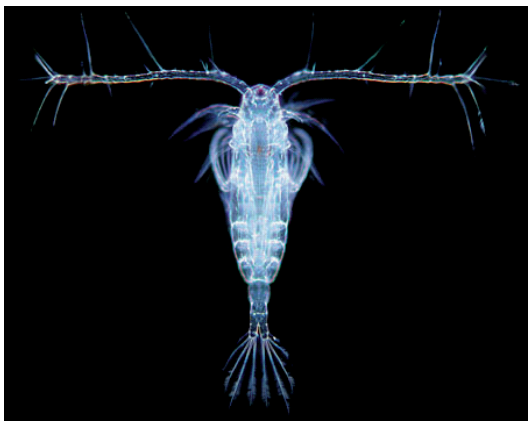


Figure 3.21: An adult copepod showing the array of sensors – setae – arrayed across the animals antenna. Similar sensors are arranged on the animals tail. Some of these sensors can detect chemical singles, while others are hydromechanically sensitive. With respect to the latter, detailed studies have shown these to be extremely sensitive to the relative velocity between the base and tip, giving neurological responses at speeds as low as 20 μms^{-1} (Yen et al., 1992).

electrode into nerve endings in copepod antennae while being stimulated by fluid motion. For a given fluid disturbance, the intensity, timing and along-antennae position of triggered setae can convey information about the size, speed and location of an approaching organism. Such information is crucial for the ensuing behavioural response (*e.g.* small ⇒ prey ⇒ attack while large ⇒ predator ⇒ escape). Understanding the nature of fluid disturbances, how they are generated and transmitted, and how they are detected, is thus central to understanding how plankton interact with each other.

3.5.1 Sensing relative velocity

Proceeding on the assumption that an organisms is equipped with setae, the bending of which elicit a neurological response, detectable hydromechanical signals arise when there is a velocity difference between the body of the organism and the fluid in its immediate surroundings. An organism imbedded in a general flow field can detect relative motion through three passive actions; it can potentially detect deformation, rotation and acceleration depending on its own physical characteristics and those of the flow field.

Fluid Deformation: A neutrally buoyant organism (*i.e.* one with the same net density as its surrounding fluid) will move pretty much at the same velocity as the fluid it displaces. This can be seen from the simplified equation of motion for a spherical particle (Eq. 3.2.24) where, for neutral buoyancy, $\alpha = 1$. Any initial velocity difference between the organism and the fluid will rapidly damp out with a time scale (a^2/μ where a is the size of the organisms), so that $\mathbf{v} = \mathbf{u}$. However, the organism with a rigid body such as a copepod, cannot conform to local gradients in the flow field over its whole body. Somewhere there will be velocity differences. Specifically, if $\mathbf{e}$ is the local rate of deformation (see Sec. 3.1.2), the velocity

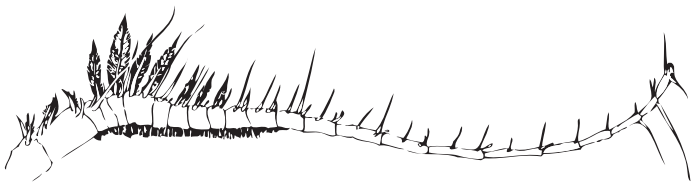


Figure 3.22: The antenna of the copeopod *Labidocera madurae*.

difference at any given setae is

$$\boldsymbol{\delta}_s = \mathbf{e}\mathbf{r}_s \tag{3.5.1}$$

where $\mathbf{e}$ is the rate of strain tensor, and $\mathbf{r}_s$ is the location of a seta on the surface of the body. This says something about the rate at which setal tips are bent due to local fluid motion. The same general rule also applies for zooplankton with "elastic" deformable bodies such as ciliates. Here the means of detection may be different (stress sensors in the cell membrane for instance so that $\boldsymbol{\delta}_s$ gives the rate at which a stress sensor is strained).

The velocity difference $\boldsymbol{\delta}_s$ is a vector, depending on the relative orientation of the organism, where its sensors are located, and the rate of strain of the fluid. While there may be extra information this relative orientation carries, I will for simplicity limit the discussion to the scale of the potential velocity difference. That is

$$\delta = a\mathbf{e} \tag{3.5.2}$$

where a is the size of the detecting organisms, and so the distance between its centre of mass and a seta on its surface.

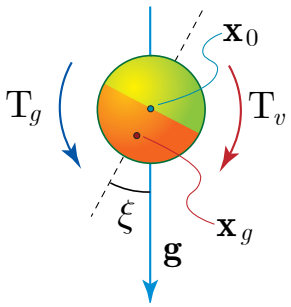


Figure 3.23: A particle with an eccentric centre of mass, entrained into a current with a horizontal component of vorticity, will come under the influence of two torques: T_g due to gravity and T_v due to frictional drag arising from relative rotation with respect to the fluid. Assuming the particle to be spherical with radius b, these are given by

$$T_g = \frac{4}{3}\pi\rho b^3 g d \sin\xi$$

and

$$T_v = 8\pi\mu b^3 \Omega^*$$

where d is the offset of the centre of mass $\mathbf{x}_g$ from the geometric centre $\mathbf{x}_0$ of the particle, g is the acceleration due to gravity, μ the dynamic viscosity, and Ω^* is the relative rotation of the particle with respect to the fluid. Since the gravitational torque is dependent on ξ, the orientation of the particle with respect to the vertical, the particle can assume a steady orientation such that the gravitational and frictional torques balance.

Eccentric mass distribution: An organism can be neutrally buoyant, but its internal mass distribution may be eccentric. That is, may be bottom heavy, its centre of mass ($\mathbf{x}_g$, where gravity acts) being off set from its geometric centre ($\mathbf{x}_0$, where the hydrostatic pressure gradient acts). At equilibrium, the organism floats with $\mathbf{x}_g$ vertically below $\mathbf{x}_0$. In a rotating flow however, friction will tend to rotate the organism up to a point where the frictional torque and gravitational torque balance. If the flow has a local horizontal component of rotation Ω (note that $\Omega = \omega/2$, where ω the vorticity) a balance between torque due to drag and torque due to gravity can be achieved when $\sin\xi = 6\eta\Omega/gd$ (Fig. 3.23), where d is the offset between the centre of mass and the geometric centre (*i.e.* $d = |\mathbf{x}_g - \mathbf{x}_0|$). However, when $\Omega > \Omega_{\max} = dg/6\eta$, no balance can be maintained, and the organisms tumbles continually.

Thus $\Omega_{\max}$ sets an upper limit of the relative rotation, and the velocity that an organism can sense is thus

$$\delta = a \min(\Omega, \Omega_{\max}). \tag{3.5.3}$$

For an elastic deformable body with an eccentric centre of mass, rotation can likewise stretch strain sensors on the "up-lifted" side of its body, while compressing those on the opposite side.

Density Contrast: Finally, an organism with a net density different from that of the surrounding fluid, will experience some "slip" in an accelerating flow (Sec. 3.2.2). Specifically, the forces that would have acted on the displaced fluid and cause it to accelerate, now act on the organism, and, because it has a different mass, cause it to accelerate at a different rate. If it is less dense that the fluid, it will accelerate faster, while it will lag behind if it is denser. Examining Eq. 3.2.24, the "slip" velocity is determined form a balance between the excess accelerative force and the drag, in a manner exactly analogous to the way sinking velocity was determined. That is, if the fluid acceleration is constant, the slip velocity after some time has elapsed will tend towards

$$\delta_{\text{acc}} \to (1-\alpha)\frac{d\mathbf{u}}{dt} = \frac{2}{9}\frac{a^2}{\mu}\frac{\rho_p - \rho}{\rho}\frac{d\mathbf{u}}{dt} \tag{3.5.4}$$

There is no direct analogy for this type of sensing in protists. However, if an organism contained an organelle with a density different from that of the cells internal fluid, the organelle would accelerate at a rate different from the rest of the organism much like the boy and his balloon in the aeroplane. Stress sensors connecting the organelle to the body could potentially measure acceleration. There is no evidence that such a mechanism exists.

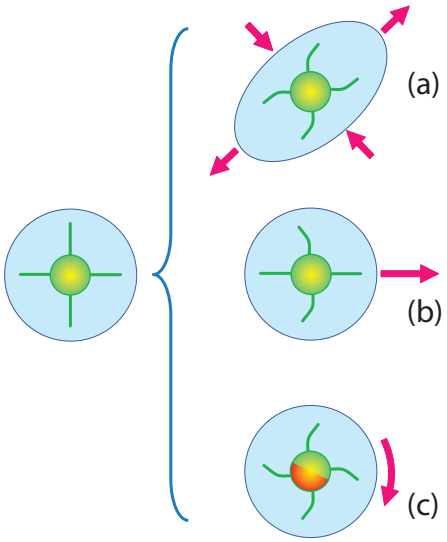


Figure 3.24: It is interesting to note, that the pattern of setal bending for each of these flow components has a different symmetry. The pattern of setal bending around the body of an organisms for the three cases (a) deformation, (b) acceleration (*i.e.* slip velocity) and (c) rotation for a bottom-heavy organism. If these flow components can be detected, the organism could potentially use these patterns to give a higher level of information.

3.5.2 Escape reactions

One of the most striking and reproducible examples of rheotactic behaviour is the escape response elicited by exposing copepods and ciliates to a siphon flow. Organisms placed in

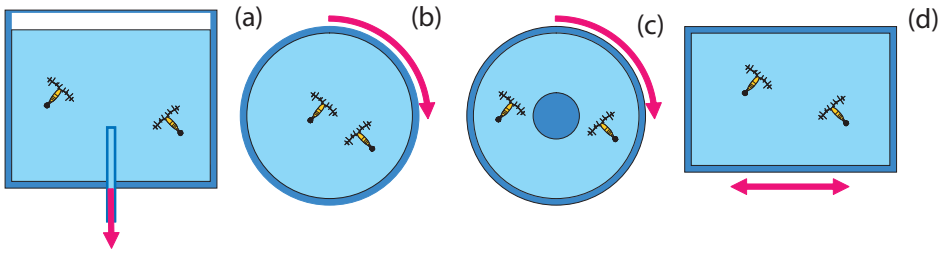


Figure 3.25: Four experimental setups: a siphon flow, a rotating cylinder, a couette flow, and an oscillating chamber

an aquarium that is slowly draining through a tube, will be inexorably be sucked towards the end of the tube. At a certain distance from the tube opening, where presumably the fluid disturbance exceeds a critical value, the organisms start to display escape behaviour (jumps or rapid swimming). But what exactly is the nature of the fluid disturbance they are reacting to. A siphon flow has 2 potential characteristics that can produce a velocity difference between an organism and the fluid, namely deformation (or rate of strain) and acceleration. Likewise, other types of flow may have other characteristics (*e.g.* shear or vorticity) that can produce similar effects. In an attempt to isolate these flow components and their effects on copepod escape responses, we (Kiørboe et al., 1999) ran a series of experiments each with different flow characteristics (Fig. 3.25): namely siphon flow (deformation and acceleration), oscillating chamber (acceleration only), couette flow (shear (*i.e.* deformation and vorticity)) and a rotating chamber (vorticity only). By quantifying the different flow components for these devices, and observing where and when copepods exhibit escape reactions with respect to these components, a means of establishing the behavioural sensitivity of these animals to fluid disturbances in nature can be established.

Siphon Flow: An idealized siphon flow is characterized by a flow rate Q though a sink. The structure of the flow can be got from simple kinematics. Continuity demands that the flux through concentric spherical surfaces of various radii r, with the siphon tip at their centre, must be equal to each other and equal to Q. That is $4\pi r^2 u_r = Q$; the radially directed fluid flow u_r times the surface area of a sphere equals Q. An organism entrained in such a flow will experience fluid deformation and acceleration.

$$\delta_{\text{def}} = \frac{aQ}{2\pi r^3} \tag{3.5.5}$$

$$\delta_{\text{acc}} = (1-\alpha) u_r \frac{\partial u_r}{\partial r} = (1-\alpha)\frac{Q^2}{8\pi^2 r^5} \tag{3.5.6}$$

As animals are transported towards the siphon tip, these signal strengths increase slowly at first, and then rapidly. By measuring the distance at which escape responses are observed, a relatively precise estimate of the threshold signal can be determined.

Oscillating Chamber: Animals enclosed in a fluid-filled oscillating chamber with no air pockets experience acceleration only. If the chamber as a whole oscillates with horizontal displacement $x(t) = X\cos\varpi t$, then all of the fluid elements within the chamber oscillate in the same manner. The velocity and acceleration of the fluid motion are given by:

$$\frac{du}{dt} = \frac{d^2x}{dt^2} = -X\varpi^2 \cos\varpi t \tag{3.5.7}$$

If the organisms have a density contrast α with respect to the fluid, then they will experience a slip velocity

$$\delta_{\text{acc}} = X(1-\alpha)\frac{\varpi^2}{\sqrt{\lambda^2+\varpi^2}} \tag{3.5.8}$$

where $\lambda = 3\eta/a^2$ is the inverse of the Stokes time scale.

Rotation: A fluid enclosed in a rotating cylinder will eventually assume a solid body rotation at the same rate as the cylinder. An organism placed in the rotating cylinder will experience the same rotation rate with respect to the vertical as the cylinder. In addition, the organism also experiences centripetal acceleration; if denser than the fluid it will be flung outwards and if less dense it will migrate to the centre of the cylinder. This of course in the principal used in centrifuges to concentrate cells. For a cylinder rotating at rate Ω, the potential signals are

$$\delta_{\text{acc}} = \frac{2}{9}\frac{a^2}{\mu}(1-\alpha)r^2\Omega \tag{3.5.9}$$

$$\delta_{\text{rot}} = a\Omega \tag{3.5.10}$$

Couette flow: The flow between two concentric cylinders (radius R_o and R_i respectively) rotating at different rates (Ω_o, Ω_i) is a couette flow. At steady state $(Du/Dt = 0)$ and in the absence of other forces, $(f = 0)$, the Navier-Stokes' equations governing the dynamics conform to

$$0 = \mu\left(\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} - \frac{v}{r^2}\right) \tag{3.5.11}$$

$$\frac{\partial p}{\partial r} = \rho\frac{v^2}{r} \tag{3.5.12}$$

These are presented in cylindrical polar coordinates (*cf.* Batchelor, 1959). The general solution to Eq. 3.5.11 is

$$v = C_1 r + \frac{C_2}{r} \tag{3.5.13}$$

(Taylor 1923). Applying the boundary conditions that the flow conforms to the rotation rates at the inner and outer cylinders, *i.e.*

$$v = R_o\Omega_o \quad \text{at} \quad r = R_o \tag{3.5.14}$$

$$v = R_i\Omega_i \quad \text{at} \quad r = R_i \tag{3.5.15}$$

gives

$$C_1 = \frac{R_o^2\Omega_o - R_i^2\Omega_i}{R_o^2 - R_i^2} \tag{3.5.16}$$

$$C_2 = \frac{R_o^2 R_i^2(\Omega_o - \Omega_i)}{R_o^2 - R_i^2} \tag{3.5.17}$$

from which the general solution follows. Settling $\Omega_o = \Omega$ and $\Omega_i = 0$, the, various components of the flow of potential influence are

$$\delta_{\text{def}} = a\Omega \frac{R_o^2 R_i^2}{r^2 \left(R_o^2 - R_i^2\right)} \tag{3.5.18}$$

$$\delta_{\text{acc}} = -\frac{2}{9}\frac{a^2}{\mu}(1-\alpha) r^2 \Omega \frac{r - R_i}{r^2 \left(R_o - R_i\right)} \tag{3.5.19}$$

$$\delta_{\text{rot}} = a\Omega \frac{r - R_i}{r^2 \left(R_o - R_i\right)} \tag{3.5.20}$$

Sensitivity: Laboratory experiments exposing *Acartia tonsa* (nauplii, copepodites and adults) to the four types of flow, show that escape responses are primarily elicited in response to fluid deformation. This can be deduced from the difference in escape responses, where responses were observed in the siphon flow (deformation and acceleration) but no response were observed in the oscillating chamber, even when the acceleration associated with the oscillations were many times higher than gravity. Copepods are very nearly neutrally buoyant (at least those tested here) so that the slip velocity is small. The threshold rate of strain to trigger escape reactions in animals was stage dependent, being higher for smaller stages (1 to 2.5 s^{-1} for nauplii), and lower for larger stages (0.5 s^{-1} for adults). Taking the size of these stages into account, the threshold velocity appeared the same about 0.03 ± 0.002 cm s^{-1} for all stages. For copepods in general, escape responses tend to be elicited at deformation rates of between 0.4 to 6 s^{-1}, while threshold velocities vary over 2 orders of magnitude.

3.6 Hydromechanical signals and remote detection

From a zooplankter's point of view, the marine environment is sparsely populated. Remote detection is thus a critical factor in how such organisms find mates, detect and capture prey, and sense and escape from predators. Chemical and hydromechanical signals are the chief conduits of information to and from plankton. While chemical signals will given some consideration in Sec. 4.7, here I would like to devote some pages to the hydromechanical signals in remote detection – how far such signals propagate, can be detected, and the information they carry.

I stress that my aim here is to uncover general rules. Specific knowledge of small scale hydrodynamics is generally advanced enough that elaborate numerical simulations can be constructed for nearly any planktonic interaction – complex body shapes, means of applying thrust, detector geometries and 3D trajectories can in principal all yield to *in silica* simulation (*e.g.* Jiang et al., 2002; Jiang and Osborn, 2004). However, the more elaborate the simulation the more difficult it is to draw general principals form the results. With this in mind, I wish to outline here a broad application of principals rather than details.

How an organism moves through water, depends on the forces that act on it. Primarily these are buoyancy, thrust and drag. Of these, thrust and drag act not only on the organism, but are also mirrored on the fluid (equal and opposite in accordance with Newton's 3^{rd} law), producing hydromechanical disturbances that spread out from the moving organism. In general, the strength of the hydromechanical disturbance depends on the size and speed of the moving organism. Here I will try to promote two ideas. Firstly that the spatial pattern of the hydromechanical signal generated by a moving organism is different for different classes of movement (*e.g.* passive sinking, uniform swimming, rapid jump). Secondly, that the spatial scale, together with a plankter's sensitivity can be used to determine its detection distance

to an organisms generating the fluid disturbance (*e.g.* potential prey, mate or predator) – detection distance being central in determining encounter rates. Further, I want to promote the idea that imbedded in the signal is higher order information: the patterns and timing of signals across a sensory array give quantitative information that a searcher can use to estimate the speed, size, location and direction of movement of a target organism. In this way, I seek to quantify the potential transfer of information to and from moving organisms through their fluid environment.

The specific hydromechanical disturbance generated by a plankter depends on the sum of the forces it exerts on the fluid, their spatial distribution, the boundary conditions imposed by the solid body of the plankter, and Re, the Reynolds number regime (*cf.* Sec. 3.1.4) in which the flow is generated. Re gives the ratio of the inertial to viscous forces operating on the fluid, and it is noteworthy that Re covers many decades from $\mathrm{O}(10^3)$ for a copepod escape jump (inertia dominates), to $\mathrm{O}(10^{-3})$ for a phytoplankton cell imbedded in a copepod feeding current (viscosity dominates) (*e.g.* Yen 2000). This large range in Re means that different dynamical descriptions must be considered; from Stokes' to the Eulerian regimes.

The first limit concerns Stokes' (or creeping) flow where exact analytical solutions can often be found for flow induced by simple force distributions and body shapes. Furthermore, because the dynamics governing Stokes' flow are linear, methods can be easily adapted for more complex systems by adding solutions together. Formally, however, Stokes' flow is valid only for very low Reynolds number. It is somewhat surprising then that for 3 dimensional bodies, Stokes' flow solutions compare favourably, at least in bulk characteristics such as drag, with observations up to $Re = 1$, with appreciable deviations only appearing at $Re > 10$ (Batchelor, 1967). Furthermore, where more exact solutions at finite Re have been found, the Stokes' flow solution generally appears as a first order approximation.

At the other extreme ($Re > 10$) where inertia becomes important, laminar boundary layer theory can be applied. This is widely used in aerodynamics and assumes that viscosity is only important in a thin laminar layer immediately adjacent to the translating body. Beyond this, the flow is inertial, and is governed only by continuity and the conservation of momentum. This type of flow has a strong fore-aft asymmetry. Momentum and vorticity generated in the laminar boundary layer are shed down-stream of the moving body, and the wake becomes turbulent at about $Re > 100$. Forward of the moving body however, the flow is undisturbed and remains laminar at these and even higher Reynolds numbers. This can be seen in photos of the flow up-stream from a sphere (Batchelor, 1967). In many instances, it is the characteristics of the up stream-flow (*e.g.* the bow-wake of an approaching fish) that is important for plankton interactions.

The flow field associated with any translating 3D body will always have characteristics of both the Stokes and Eulerian regime. Sufficiently close to the body, viscosity will dominate, and sufficiently far away, viscosity becomes less important and inertia tends to dominate. The distance at which this transition takes place is the boundary layer thickness; $d = aRe^{-1/2}$ where a is the size of the translating body. As a rule of thumb, for reaction events outside of the wake, a comparison of the reaction distance R, with the boundary layer thickness gives a good indication of the relevant regime: $R < d$ for Stokes' and $R > d$ for Eulerian.

This section is primarily concerned with the scale of the hydromechanical signal generated by different motile behaviours of target plankton and scanning modes (*e.g.* feeding currents) of searcher organisms. Specifically, given known mechano-sensitivities for either neurological or behavioural responses, the feasibility of mechanoreception as a means of remote detection

for differing behavioural classes is explored. The specific types of behaviour and processes considered here include

- passive particle in a deforming flow, *e.g.* a phytoplankton cell entrained in a feeding current or approached by a free-swimming copepod.
- eccentric (*i.e.* bottom heavy) particle in a rotating flow *e.g.* a nauplius entrained in a feeding current.
- sinking particle, *e.g.* faecal pellet or marine snow detected by a flux feeding copepod (*e.g. Oithona*)
- feeding current, *e.g.* as produced by calanoid copepods, ciliates and bivalve larvae
- constant swimming at low Re, *e.g.* ciliates, nauplii, rotifers detected by ambush feeding copepod
- beating feeding appendages and hydro acoustics
- constant swimming at high Re, *e.g.* fish larvae
- rapid jump, *e.g.* copepod escape reaction

These are arranged more or less in order of increasing Re.

3.6.1 Sinking and swimming – stokeslet and stresslet

Two of the most generic flow fields in plankton ecology are those associated with sinking particles and self-propelled organisms at low Reynolds number. In Sec. 3.3 it was noted that the flow associated with any physical situation can in principle be deduced from a multi-pole expansion, where certain far field characteristics of the flow are dictated by the gross force distribution exerted by a moving body on the fluid. In particular, a single force acting on the fluid results in a stokeslet, while two equal and opposite forces separated by a small spatial distance generate a stresslet. The stresslet attenuates faster in space (as r^{-2} with distance from the source) compared to the stokeslet (as r^{-1} with distance from the source). Here we come to an immediate distinction between the flow generated by a sinking body and that generated by a self-propelled body, even when the bodies in question are the same size and shape, and moving at the same speed through the fluid.

At steady state, when the body is moving at a constant speed, the force balance on the *body* is zero; the propulsive force – buoyancy for a sinking body, thrust for a self-propelled body – exactly balances drag. However, the force balance on the *fluid* is not zero in both cases. While thrust and drag have mirror forces acting on the fluid, buoyancy is a body force with no corresponding equal and opposite force acting on the fluid. The consequence is that the lowest order multipole describing the flow generated by a sinking body is a stokeslet, while that generated by a self-propelled body is a stresslet.

3.6.2 Passive sinking body

The detection, interception and ingestion of settling particles such as marine snow and faecal pellets by copepods (*e.g. Oithona* spp.) and other zooplankton, is thought to be an important factor controlling the net vertical flux of carbon in the open ocean (*e.g.* Jackson 1993, González & Smetacek 1994). The interception of these sinking particles is in general size dependent – large particle sink faster, generating a stronger signal that can be detected further away by intercepting zooplankters.

While the flow associated with a sinking spherical particle has been considered (Sec. 3.2.1), it can also be derived fairly easily from the superposition of some fundamental multipoles.

For a sphere of radius a, sinking with a velocity $-U\hat{\mathbf{z}}$ ($\hat{\mathbf{z}}$ is the unit vector positive upwards) the drag exerted by the particle on the fluid is given by $\mathbf{f}_d = -6\pi a\mu U\hat{\mathbf{z}}$ which gives the strength and orientation for the stokeslet component. The second element of the flow field concerns the boundary conditions, *i.e.* that there is no flow and no net stress on the body's surface.

For a translating spherical body, it turns out that a potential dipole acting at the geometric centre of the body is exactly the entity that is required. Finally, the stokeslet + potential dipole flow, is relative to the flow at large distances. In order for the flow to be written relative to the settling body, we add on the relative far field motion of the fluid. The final result (*cf.* Sec. 3.2.1) is

$$\mathbf{u} = U\begin{bmatrix} -\cos\theta \\ \sin\theta \end{bmatrix} + \frac{3}{4}\frac{aU}{r}\begin{bmatrix} 2\cos\theta \\ -\sin\theta \end{bmatrix} - \frac{1}{4}\frac{a^3U}{r^3}\begin{bmatrix} 2\cos\theta \\ \sin\theta \end{bmatrix} \tag{3.6.1}$$

(The notation for these and following vector notations is $\mathbf{u} = [u_r, u_\theta]$, u_r being the radial and u_θ the zenith, or polar components of the flow field.)

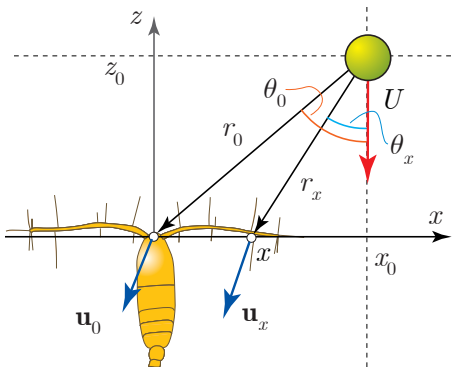


Figure 3.26: The detection of a sinking particle (swimming speed U, size a by a stationary raptorial predator (antenna length b). Detection occurs when the magnitude of the velocity difference between the fluid and the body at a setae located at x exceeds a critical threshold.

Now let us return to the particle interception problem. Consider an ambush feeding copepod, with antenna length b, feeding on settling particles of radius a that sinking with speed U. As a sinking body falls past the copepod, it can, in general, cause the copepod to move along with the ambient fluid. This is particularly clear if the particle is large with respect to the copepod. The velocity of the copepod (body + antennae) induced by the settling particle is given by

$$u_0 = \frac{3}{4}\frac{aU}{r_0}\begin{bmatrix} 2\cos\theta_0 \\ -\sin\theta_0 \end{bmatrix} - \frac{1}{4}\frac{a^3U}{r_0^3}\begin{bmatrix} 2\cos\theta_0 \\ \sin\theta_0 \end{bmatrix} \tag{3.6.2}$$

where (r_0, θ_0) is the position of the centre of mass of the copepod relative to the sinking particle (Fig. 3.26). A seta positioned somewhat closer to the settling body at (r_x, θ_x) will be at a location where the ambient fluid velocity, $\mathbf{u}_x$, is different (for the relevant formula, simply replace (r_x, θ_x) for (r_0, θ_0) in Eq. 3.6.2). The seta will therefore experience a velocity difference $\delta\mathbf{u} = \mathbf{u}_x - \mathbf{u}_0$. Fig. 3.27 shows an example of the velocity components across a copepod's antenna as a particle sinks past.

Assuming that detection occurs when the magnitude of the velocity difference $\delta u = |\mathbf{u}_x - \mathbf{u}_0|$ exceeds a given sensitivity s, the distance $R^* = R - b$ can be estimated. (This notation will

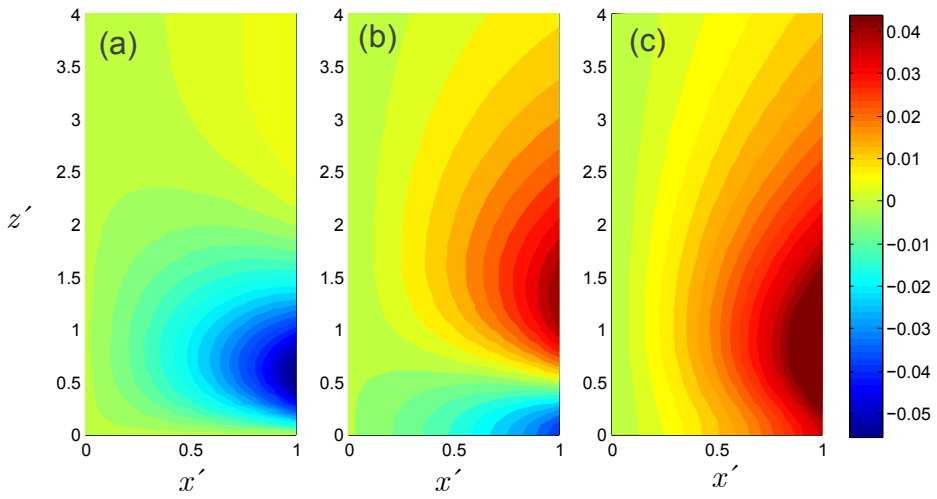


Figure 3.27: The pattern of the hydromechanical signal across an antenna (oriented horizontally) produced by a particle sinking vertically and a distance $x_0 = 2b$ to the right of the centre of mass of the copepod. The antenna and sinking track of the particle are co-planar. The signal strength is normalized by U the sinking speed, distance by b, and time by b/U. (a) is the along antenna signal component, (b) the across component, and (c) the absolute value of the signal strength. x' is the along antenna position (in units b) while z' is the vertical location of the centre of the particle. From the copepod's point of view z' is a measure of the *time* at which it experiences a particular signal pattern: $t = z/U$.

be followed: R denotes the detection distance from the centre of the detector array, while $R^* = R - b$ is the distance from the tip of the detector array.) Even for this relatively simple case, the distance R^* at which detection occurs can be a complicated function of geometry. A reasonable estimate can be deduced from the instant where the particle is closest to the copepod (*i.e.* $\theta_0 = \theta_x = \pi/2$, Fig. 3.27), the velocity difference δw, is in the vertical direction and given by

$$\delta w = \frac{3U}{4}\frac{ab}{r(r-b)} + \frac{3U}{4}\frac{a^3 b}{r^2(r-b)^2} + \frac{U}{4}\frac{a^3 b^3}{r^3(r-b)^3} \tag{3.6.3}$$

which can in principal be solved numerically for $\delta w = s$, the solution of r being the detection distance R^*. The case of most interest is

$$R^* \approx \frac{b}{2}\left(\left(1 + 3\frac{a}{b}\frac{U}{s}\right)^{1/2} + 1\right) \tag{3.6.4}$$

which is applicable for $R^* > a$, *i.e.* when the detection distance is larger than the particle size – detection distances less than this are direct contacts which hardly count as remote detection.

With regards mechanoreception, we note that the settling velocity U is dependent on the particle size a. For example, relatively compact particles such as faecal pellets settle at a rate determined by Stokes' law: $U = 2ga^2\Delta\rho/(9\mu)$ where $\Delta\rho$ is the density difference between the fluid and the particle. Firstly, this implies there is a particle size limit below which there is no detection. That is, when $R^* < a$, the particle cannot be detected which leads to a size

limit for detection

$$a_0 = \left(\frac{6\mu s}{g\Delta\rho}\right) \tag{3.6.5}$$

For *Oithona* ($b = 0.75$ mm, $s = 40$ μm/s) feeding on faecal pellets of density 1.15 g cm^{-3}, this gives $a_0 = 13$ μm, which is about half the size of an *Acartia* faecal pellet ($a = 25$ μm).

3.6.3 Self-propelled body

Plankton that swim gain from finding resources but pay a penalty in that they are hydrodynamically more conspicuous exposing themselves to the risk posed by rheotactic predators (Landry, 1980; Jonsson and Tiselius, 1990; Svensen and Kiørboe, 2000).

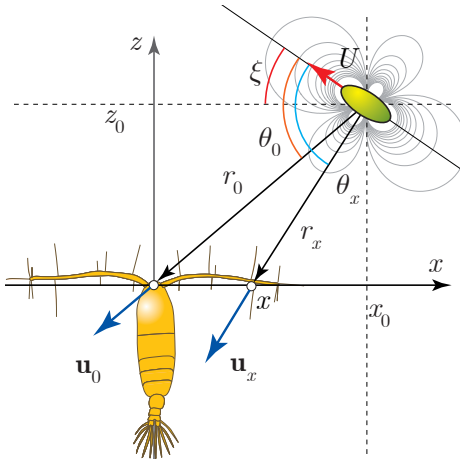


Figure 3.28: The detection of a self-propelled organism (swimming speed U, size a by a stationary raptorial predator (antenna length b) depends on the geometry of the swimming path. For a linear path, and assuming the path and detector array are co-planar, this is defined by the position (x_0, z_0) of the swimming organisms with respect to the centre of mass of the detector at a given time, an the angle ξ the path makes with a fixed direction (the horizontal here).

The net force exerted on the fluid by a neutrally-buoyant, self-propelled body moving at uniform velocity U, is zero – thrust balances viscous drag on the body – so the lowest order multipole describing the flow is a stresslet. The simplest case is a self-propelled sphere of radius a, where the thrust is applied on the surface of the sphere and acts through it's centre of mass. While the drag acts over the surface of the sphere, to first order it can be considered as an effective point force diametrically opposed to the thrust. The strength of the dipole forces are thus $f = 6\pi\mu a U$, are equal and opposite, and are separated by the distance $2a$ gives

$$u_r = \frac{3}{2}\frac{Ua^2}{r^2}(1 - 3\cos^2\theta) \tag{3.6.6}$$

$$u_\theta = 0 \tag{3.6.7}$$

(Note, a fuller flow description including translation and boundary conditions on a solid spherical body leads to the "squirmer" solution Sec. 3.4.4). In general, the velocity difference between a sensor at (r_x, θ_x) and the centre of mass of a copepod at (r_0, θ_0) (Fig. 3.28) is given by

$$\delta\mathbf{u} = \frac{3}{2}Ua^2\left(\frac{1 - 3\cos^2\theta_x}{r_x^2}\hat{\mathbf{r}}_x - \frac{1 - 3\cos^2\theta_0}{r_0^2}\hat{\mathbf{r}}_0\right) \tag{3.6.8}$$

where the vectors $\hat{\mathbf{r}}_x$ and $\hat{\mathbf{r}}_0$ are the unit radial vector evaluated at O and x respectively. Assuming that detection occurs when $|\delta\mathbf{u}| > s$, the detection distance R^*, can be deduced.

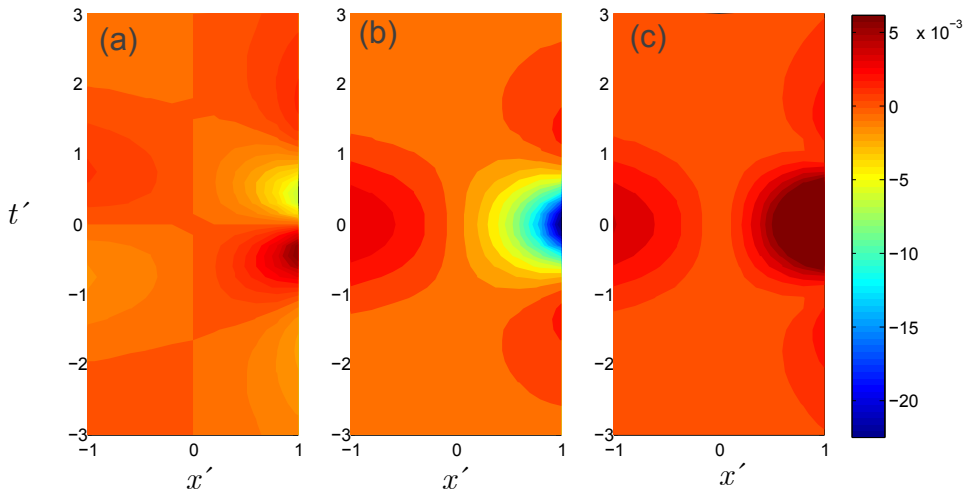


Figure 3.29: The hydromechanical signal (Eq. 3.6.8)across an antenna (oriented horizontally) produced by a self propelled organism swimming vertically ($\xi = -\pi/2$) and a distance $2b$ to the right of the centre of mass of the copepod. The signal strength is normalized by U, distance by b, and time by b/U. (a) is the along antenna signal component, (b) the across component, and (c) the absolute value of the signal strength

The general case is algebraically cumbersome. Instead, consider first the case where detection occurs close to the detector ($R^* \ll b$).

This necessarily means that $r_0 = R \gg R^*$. In this case, the velocity at $r_x = b$ dominates in Eq. 3.6.8, and the detection distance is given by

$$R^* < b: \qquad R^* = a\left(\frac{3}{2}\frac{U}{s}|3\cos^2\theta - 1|\right)^{1/2} \tag{3.6.9}$$

On the other hand, when the detection distance is large compared to the antenna size b, we have $\theta_0 \approx \theta_x (= \theta)$, and $r_0^2 = r_b^2 + 2r_b b\cos(\theta - \xi)$. The detection distance R^* is given by.

$$R^* > b: \qquad R^* = \left(3a^2 b\frac{U}{s}|3\cos^2\theta - 1|\cos(\theta - \xi)\right)^{1/3} \tag{3.6.10}$$

where ξ is the angle between the approach direction and the alignment of the antenna (Fig. 3.28). The conditions relating when these scaling rules apply can be specified by noting the magnitude of the estimates. Specifically, Eq. 3.6.9 applies when $aU^2/(bs^2) < 1$, whereas Eq. 3.6.10 applies when $aU^2/(bs^2) > 1$.

Observations of the reaction distance of the copepods feeding on motile protists can be used to test these predictions. For instance, *Acartia tonsa* detect the ciliate *Loboea strobila*(length=130 μm, width=65 μm, U = 0.5 to 1 mm/s) at a distance of about 0.5 mm (Jonsson and Tiselius, 1990). In comparison, when feeding on *Strombidium reticulatum*, (length=43 μm, width=30 μm, U = 0.5 to 1 mm/s), they observed the detection distance to be considerably less, with nearly all attacks occurring at less than 0.25 mm. Substituting these values into Eq. 3.6.9 and assuming s = 40 μm/s gives reaction distances of 0.40 to 0.56 mm and 0.13 to 0.19 mm for *Loboea strobila* and *Strombidium reticulatum* respectively. More recently, Svensen and Kiørboe (2000) reported observed reaction distances R =

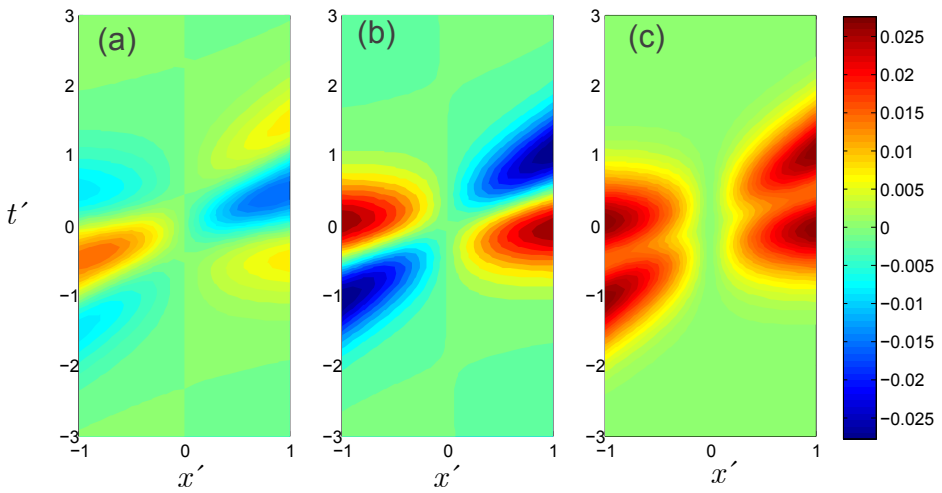


Figure 3.30: The hydromechanical signal across an antenna (oriented horizontally) produced by a self propelled organism swimming horizontally ($\xi = 0$) and a distance b above the centre of mass of the copepod. The signal strength is normalized by U, distance by b, and time by b/U. (a) is the along antenna signal component, (b) the across component, and (c) the absolute value of the signal strength

140± 70 μmfor *Oithona similis* feeding on *Gymnodinium dominans* ($a = 19$ μm, $U = 0.58$ mm/s). Again, substituting into Eq. 3.6.9 we get $R = 125$ μm. In all instances, estimated and observed detection distances correspond relatively well.

Directional, temporal and spatial information: For both the settling particle and swimming organisms, the assumed criterion for detection is simply that the absolute velocity difference at a detector $|\mathbf{u}_0 - \mathbf{u}_x|$, exceeds a detection threshold, s. While this approach certainly says something about the feasibility of detection, it says nothing about higher information content that can be used in locating the source of the hydromechanical signal. In this, we note that copepod setae appear to be hinged so as to allow velocities to be measured in different directions (Yen et al., 1992) giving possible locator ability. The most obvious case this that for locating a small swimming organism modelled as a stresslet. Because the velocity disturbance is always radial (*cf.* $\mathbf{u}_0$ and $\mathbf{u}_x$ in Fig. 3.28), if a copepod can measure the direction of the velocity disturbance at the same time at two locations along its antenna, a simple triangulation will point directly back to the swimming organism.

In addition, there is also temporal and spatial information which can be used to locate swimming organisms. Different setae positioned at different locations along the antenna receive different directional and intensity information over time depending on the location and swimming direction of the swimming organism. Compare Fig. 3.29 and Fig. 3.30 for instance.

3.6.4 Feeding currents

Many zooplankton species including copepods and ciliates have developed feeding currents. In copepods, these feeding currents are generally generated by movements of the second antennae, first maxillae and the maxillipeds or combinations there of (*e.g.* Koehl and Strickler, 1981; Strickler, 1982; Paffenhöfer et al., 1982). The general shape of the feeding current is a funnel like, inwardly directed flow forward of the animal (Strickler, 1985; Tiselius and

Jonsson, 1990).

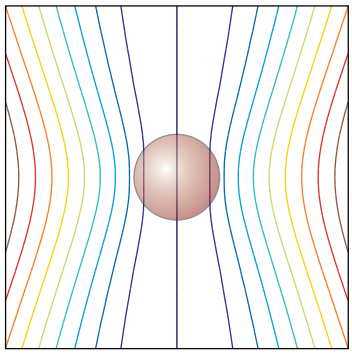


Figure 3.31: The streamfunction ψ of a spherical pump. Distance is normalized by pump radius a and flow rates by pump velocity U. The streamfunction is given by

$$\psi = -\frac{1}{2}Ur^2 \sin^2\theta \left(\frac{3a}{2r} - \frac{a^3}{2r^3}\right)$$

The flow within the pump is uniform downwards, and conforms to a stream function $\psi_0 = -\frac{1}{2}Ua^2 \sin^2\theta$.

The simplest physical model of a feeding current is that of a gravitationally tethered copepod, where the thrust of its feeding current exactly balances its negative buoyancy (Strickler, 1982). As with a sinking particle, a feeding current exerts a net force on the fluid and leads immediately to the stokeslet model of a copepod feeding current introduced by Tiselius and Jonsson (1990). The spherical pump model (Kiørboe and Visser, 1999) is a slight variation of this is where the thrust is distributed over a finite spherical volume of space, radius a through which fluid is pumped at a uniformly at speed U. The velocity field for such a flow is given by:

$$\mathbf{u} = -\frac{3}{4}\frac{aU}{r}\begin{bmatrix} 2\cos\theta \\ -\sin\theta \end{bmatrix} - \frac{1}{4}\frac{a^3U}{r^3}\begin{bmatrix} 2\cos\theta \\ \sin\theta \end{bmatrix} \tag{3.6.11}$$

and the associated stream function[14] is plotted in Fig. 3.31. Several studies confirm this general flow structure (Tiselius and Jonsson, 1990; Yen et al., 1991; Kiørboe and Saiz, 1995)

Detection of a copepod's feeding current by prey items and their subsequent escape ability are important factors driving prey selectivity in the plankton. This process has been variously investigated in many aspects (Haury et al., 1980; Fields and Yen, 1996b; Kiørboe et al., 1999). I will skip this issue for the moment, focussing instead on the inverse process; viz. on how the feeding current can alert the predator of the presence of a prey item entrained into it.

3.6.5 Feeding current as a "radar"

Copepods can apparently detect non-motile particles being swept towards them in their feeding current while they are still several particle diameters away. While chemical signals can certainly be a cue for this detection, hydromechanical signals are likely also to play a role (*e.g.* Buskey, 1984; Vanderploeg et al., 1990; DeMott and Watson, 1991; Bundy et al., 1998).

A neutrally buoyant particle in a flow field will in general accelerate at the same rate as the fluid. Only if a particle is more (less) dense, will it accelerate slower (faster) than the ambient fluid (Kiørboe and Visser, 1999). However, a neutrally-buoyant, rigid body embedded in a deforming or rotating flow field, can generate a local flow disturbance simply by its physical presence. The reason for this is that the flow must match on the boundary of the rigid particle; a condition that can potentially lead to a perturbation of the flow.

[14] Note that this flow field is simply a re-interpretation of the flow associated with a sinking sphere relative to the far-field reference frame: 2 models for the price of one.

The flow field disturbance $\mathbf{u}'$ generated by a rigid spherical body, located within a general deforming flow field is given by

$$\mathbf{u}' = -\frac{a^5}{r^5} E \cdot \mathbf{x} - \frac{5}{2}\left(\frac{a^3}{r^5} - \frac{a^5}{r^7}\right)\mathbf{x}(\mathbf{x}\cdot E\cdot\mathbf{x}) - \frac{1}{2}\frac{a^3}{r^3}\omega\times\mathbf{x} \tag{3.6.12}$$

where E is the rate of strain tensor, and ω is the flow vorticity. A special case for simple irrotational, deformation flow is

$$\mathbf{u}' = \frac{\partial u_r}{\partial r}\frac{a^3}{4r^4}\begin{bmatrix} (3\cos^2\theta - 1)(3a^2 - 5r^2) \\ 6\sin\theta\cos\theta \end{bmatrix} \tag{3.6.13}$$

This flow perturbation can contribute to the remote detection of an otherwise inert plankton cell in a copepod feeding current. Consider a copepod of effective diameter $2a$ generating a stokeslet feeding current of strength U. The rate of strain imposed on the fluid directly in front of the copepod ($\theta = 0$) is

$$e(r) = \frac{\partial u_r}{\partial r} = \frac{3}{2}\frac{aU}{r^2} \tag{3.6.14}$$

where r is the distance from the copepod. The perturbation field for this flow field in the direction of the principal axis of deformation is given by

$$u_r'(r') = \frac{5}{2} e \frac{b^3}{r'^2} \tag{3.6.15}$$

where b is the radius of the plankton cell, and r is the radial distance measured from the centre of the plankton cell. The plankton cell is detected when the perturbation velocity u_r' exceeds the threshold detection velocity s at a sensor. Approximating the distance from the plankton cell to the sensor by $r' = r - a$, the separation distance R at which detection occurs is given by:

$$R = \frac{a}{2}\left(1 + \left(1 + 2\left(\frac{15b^3U}{a^3 s}\right)^{1/2}\right)^{1/2}\right) \tag{3.6.16}$$

As an example, consider a copepod of equivalent radius $a = 0.5$ mm, generating a feeding current of magnitude $U = 10$ mm/s and sensitivity $s = 0.02$ mm/s Yen et al. (1992). A plankton cell of radius $b = 0.05$ mm could then be detected at a distance of $R = 0.89$ mm ($R^* = 0.39$ mm). Jonsson and Tiselius (1990) report observations of *Acartia tonsa* (length 800 μm) interacting with latex beads of diameter 76 μm. While the nature of this interaction appeared to be a random collision rather than an active attack, we can use this as an example to test the feasibility of remote detection of inert particles in a feeding current. The maximum feeding current speed for *A. tonsa* is 8 mm/s, dropping off to about 4 mm/s at 0.5 mm from the capture area (Jonsson and Tiselius, 1990). Taking $a = 0.2$ mm, $b = 0.038$ mm, $U = 8$ mm/s and $s = 40$ μm/s gives $R = 0.42$ mm ($R^* = 0.22$ mm). It is tempting to compare this with the distance at which *A. tonsa* "attacked" the latex beads (Jonsson and Tiselius, 1990); about 0.3 mm from the copepod's head, although the exact process involved cannot be confirmed.

3.6.6 Relative rotation

A rigid spherical body of radius b that rotates at a rate Ω with respect to a fluid, generates a flow field disturbance $\mathbf{u}'$ that can be approximated by

$$u'_\phi = -\frac{\Omega b^3}{r^2}\sin\theta \tag{3.6.17}$$

where $\theta = 0$ is the axis of rotation, and the velocity u'_ϕ is directed perpendicular to the radial direction.

While many organisms have marked rotational locomotion (bacteria, rotifers, ciliates), they apply no net torque on the fluid. However, relative rotation and a net torque can result for

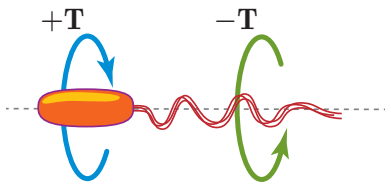


Figure 3.32: A self rotating organism such as a bacterium or a citiate applied no net torque on the fluid – the torque applied by a rotating flagellar bundle for instance is balanced by a drag torque in the opposite rotational sense. In general however, these equal and opposite torques are offset by a distance of the order of the body size of the organism leading to a rotational couplet – a quadrapole – force distribution acting on the fluid.

a particle with an eccentric centre of mass interacting with ambient fluid rotation (Kessler, 1986). Eccentric bodies entrained into a flow with a component of vorticity can in principle generate a rotlet perturbation of the flow.

A feeding current for instance, generally has inherent vorticity, ω, which for a stokeslet model gives

$$\omega = -\frac{3aU}{2r^2}\sin\theta \tag{3.6.18}$$

(Visser and Jonsson, 2000) where U is the feeding current speed, and a is the radial dimension of the pumping zone. This is effectively taken as half the copepod body width.

A particle with an eccentric centre of mass, entrained into a current with a horizontal component of vorticity, will come under the influence of two torques: T_g due to gravity and T_v due to frictional drag arising from relative rotation with respect to the fluid (Fig. 3.23). However, there is a limit to this balance giving a maximum relative rotation of $\Omega^*_{\mathbf{max}} = dg/(6\eta)$. That is, if the absolute fluid rotation is greater than Ω^* then the particle begins to tumble so that the relative rotations does not exceed is Ω^*. Thus, given a sensitivity s, the detection distance for a biased particle in a feeding current has two forms depending on ω, the vorticity of the feeding current, and the maximum particle rotation rate. Specifically

$$\omega < 2\Omega^*_{\mathbf{max}} \quad : \qquad R^*(R^*+a) = \sqrt{\frac{3Uab}{4s}\sin\theta} \tag{3.6.19}$$

$$\omega > 2\Omega^*_{\mathbf{max}} \quad : \qquad R^* = \sqrt{\frac{\Omega^*_{\mathbf{max}} b^3}{s}} \tag{3.6.20}$$

Estimates of the separation distance, d, appear to be of the order of 1% of the radius of the organism (Fields and Yen, 1997). This suggests a size dependence on $\Omega^*_{\mathbf{max}} = \gamma b$, where

$\gamma \approx 90$ cm^{-1} s^{-1}. For a typical copepod feeding current, (*e.g.* $a = 0.2$ mm, $U = 10$ mm/s), and their prey (b <0.05 mm), flow field rotation nearly always exceeds $\Omega^*_{\mathbf{max}}$ so that Eq. 3.6.20 applies. That is $R^* \approx b^2(\gamma/s)^{1/2}$. Thus, eccentric particles with radius $b < (s/\gamma)^{1/2}$ cannot be detected by their associated vorticity field. For $s = 40$ μm/s, this gives $b_{\min} = 66$ μm. Particles larger than this, can in principal be remotely detected. For instance $b = 100$ μm, $s = 40$ μm/s gives $R^* = 150$ μm. Note that any flow disturbance generated by this mechanism will also have an associated deformation flow as discussed in Sec. 3.6.5. Given the analysis here, detection of the deformation field will in all likelihood occur at a greater distance than detection of the rotational field component.

3.6.7 Passive particles approached by a self-propelled predator.

Completing the examination of the detection of inert particles by hydromechanical signals, consider one final case, namely that of the remote detection by a self propelled organism of a non-motile particle in its vicinity (Bundy et al 1998). At low Reynolds number, a self-propelled organism such as a cruising copepod, generates a predominantly stresslet flow field which has inherent deformation and vorticity. Much as for the case of a feeding current, an inert solid particle within a stresslet flow field generates a perturbation flow that can, in principle, be remotely detected. In order to explore the feasibility of this mechanism, I compute the rate of strain in front of a self-propelled organism given by

$$e(r) \approx \frac{\partial u_r}{\partial r} = 6\frac{a^2 U}{r^3} \tag{3.6.21}$$

Substituting into Eq. 3.6.14 and for a given sensitivity s, detection occurs at the distance R given by the solution of

$$R^3(R-a)^2 = 15a^2b^3\frac{U}{s} \tag{3.6.22}$$

Observations reported by Bundy et al (1998) of *D. sicilis* (length 1.2 mm) attacking polystyrene beads (diameter 50 μm) appear to follow this scenario. The copepod swims slowly towards the bead at a speed of 1 to 2 mm/s ($Re = 0.5$ to 1). As the copepod approaches, the bead begins to move in advance of the copepod. Soon after, at about 1 mm distance, the copepod attacks the bead. Setting $a = 0.5$ mm, $b = 0.025$ mm, $U = 2$ mm/s and $s = 0.02$ mm/s gives $R = 0.65$ mm. It appears feasible therefore that inert particles can be remotely detected by moving organisms due to a perturbation to their own flow field. That is, as first suggested by Zaret (1980); being non-motile is no guarantee of being hydrodynamically inconspicuous.

3.6.8 Beating appendages and acoustic signals

Copepods and other zooplankton use beating appendages in propelling themselves through the water, or in generating feeding currents. The oscillatory signals associated with this behaviour have often been suggested as a means by which copepods can remotely detect these organisms (Kirk 1985). While copepods do not appear to be sensitive to pressure per se (Lenz & Yen 1993, Bundy et al 1998), they are sensitive to the velocity fluctuations associated with the travelling sound wave. Indeed, Lenz & Hartline (1999) used vibrational stimuli to directly elicit remote escape/attack responses in a calanoid copepod.

The signal transmitted to the fluid from beating appendages can be modelled as an acoustic dipole (Lighthill, 1978). The strength of the velocity fluctuation attenuates differently $\mathbf{u}$ in space depending on the distance from the source r compared to the acoustic wave-length

$\lambda = 2\pi/k$. For an appendage of radial dimension a, oscillating at a frequency ϖ (radians / sec) with a peak velocity U, the amplitude of the velocity fluctuations are given by $u \approx Ua^3/r^3$ for $kr < 1$, and $u \approx Uk^2a^3/r$ for $kr > 1$. Here k is the wave number defined as $k = \varpi/c$, and c is the speed of sound in water. For the marine environment, c lies between 1450 and 1550 m/s.

For the typical frequencies of beating appendages, 10 to 1000 Hz, the relatively high speed of sound in seawater means that the wavelength of the associated acoustic wave is 1.5 to 150 m; much greater than the distances relevant for an individual plankton. Thus for the detection of beating appendages, the relevant dynamics are described by near-field acoustics(Visser, 2001). This has two important consequences. Firstly, the attenuation of the signal is nearly always rapid, falling off with a r^{-3} power law with distance from the source. Secondly, because the wavelength is long, the velocity field itself will serve only to translate the detector (copepod) as a whole yielding no velocity difference between the fluid and the detector. The detectable signal, u', is provided by the gradient of the velocity field and the size of the detector b. That is $u' = 3Ua^3br^{-4}$. This is a fundamental difference in the detection of acoustic signals by plankton versus benthic organisms, where being fixed in space means that absolute velocity is relevant.

As pointed by Yen and Strickler (1996)), the effect of the beating appendages in suspension feeding copepods is to create both an oscillating component and a mean feeding current. The same can be said for free swimming copepods, with the mean component being their motion through the fluid. In each case, the mean component is associated with a velocity field that attenuates less rapidly (r^{-1} for feeding current and r^{-2} for self-propelled) than the near field acoustic signal (r^{-4}). This would suggest that signals associated with vibrations are of minor importance.

When could sound be important?: The foregoing analysis assumed that the whole body of the copepod was essentially neutrally buoyant. However, if the antennae were of a significantly different density from the rest of the copepod's body, a "slip" velocity could be established between the fluid and setae.

3.6.9 High Reynolds number swimming.

While plankton generally inhabit a low Reynolds number regime, from time to time they make excursions into the high Reynolds number (Eulerian) regime. Hydrodynamic signals generated by high Re motion differ from those at low Re primarily in the volume of fluid influenced by viscosity. At low Re, this volume is essentially infinite. At high Re, the influence of viscosity is limited to a laminar boundary layer surrounding the moving body. High Re motion is characterized by a strong fore-aft asymmetry. Forward of the moving body, the fluid is largely unaffected and is only driven by its own inertia and continuity. Aft of the moving body however, vorticity and momentum generated at the body's surface are shed to the bulk of the fluid. Thus, at high Re, the hydromechanical signals associated with an approaching or retreating organism are different.

Immediately adjacent to a translating body, is the laminar boundary layer wherein momentum and vorticity are imparted from the body to the fluid by viscosity. This boundary layer has a thickness of $O(aRe^{-1/2})$ and is a region of intense momentum and vorticity transmission to the fluid that is advected aft of the translating body. Within the wake, momentum and vorticity are transmitted to the bulk of the fluid. This can be seen as a zone where the fluid relaxes from the disturbance generated by the translating body and can be modelled as a translating

point source of momentum (*cf.* Batchelor, 1967). Beyond the region where viscosity plays an important role, the flow is essentially inviscid and irrotational for an otherwise undisturbed fluid. That is, conservative properties of the flow, such as momentum and vorticity, are only transmitted by advection. This is potential flow.

3.6.10 Approach and attack of fish larvae

The ability of planktivorous fish to successfully approach, attack and capture rheotactic copepod prey, is dependent on the final approach speed of the fish prior to its attack lunge. Too fast and the prey is alerted and escapes before an attack is made (Viitasalo et al., 1998). The fluid disturbance forward of a swimming fish can be modelled as that associated with a translating sphere at high Re where the forward hemisphere approximates the snout of the fish. Beyond the laminar boundary layer, the flow field is non-turbulent and described by a potential dipole

$$\mathbf{u} = U \begin{bmatrix} -\cos\theta \\ \sin\theta \end{bmatrix} + \frac{U}{2}\frac{a^3}{r^3} \begin{bmatrix} 2\cos\theta \\ \sin\theta \end{bmatrix} \tag{3.6.23}$$

Directly in front of the approaching fish ($\theta = 0$) the deformation rate of the potential flow is given by

$$e(r) = \frac{\partial u_r}{\partial r} = 3U\frac{a^3}{r^4}$$

giving a detection distance

$$R = \left(3U\frac{a^3}{e^*}\right)^{1/2}$$

where $e^* \approx s/b$ is the critical deformation rate. In a series of experiments (Viitasalo et al., 1998), the rheotactic copepod *Eurytemora affinis* were seen to successfully escape stickleback *Gasteroseus aculeatus* when their approach speed exceeded about 1 cm/s, where escapes took place at a distance of about 0.3 cm. The rheotactic ability of *Eurytemora affinis* was quantified in siphon experiments giving critical deformation rates e^* of about 2 s^{-1} ($s \approx$ 1 mm s^{-1}). This rheotactic ability suggests the copepod could detect an approaching fish (potential dipole, $a = 1$ mm, $U = 1$cm/s) 0.3 mm – in accordance with the observations.

3.6.11 Vortex rings and escape jumps

Copepods, when alerted to proximate danger, can perform spectacular escape runs, either in the form of high velocity darts, constant swimming, or a sequence of jumps (*e.g.* Yen and Strickler, 1996). Escape speeds can reach up to 500 mm/s, corresponding up to 200 body lengths per second. That these events are inherently high Reynolds number can be seen in Schlieren image of the hydrodynamic "footprint" left in the wake of an escaping *Euchaeta rimana* (Yen and Strickler, 1996), showing mushroom-shaped vortices shed at about 35 ms intervals and 5 mm separation. The average Reynolds number for this behaviour is $Re =$ 100, although it may be considerably higher at the instant that each vortex is shed.

While powerful escape jumps are useful in removing the copepod from immediate risk, they also generate strong signals that can alert the predator to their presence. Indeed, some predator attacks appear to be triggered by the prey escape reaction (Yen and Strickler, 1996). Artificial vortices have in fact, been used in the past to elicit responses in zooplankters, *e.g.*

in examining the mate locating abilities of the copepod *T. longicornis* (Yen et al., 1998), and in the attack response of chaetognaths (Nishii 1998). The morel here is clear: if a prey is to escape, it better make it count since a half-hearted attempt is worse that no attempt at all.

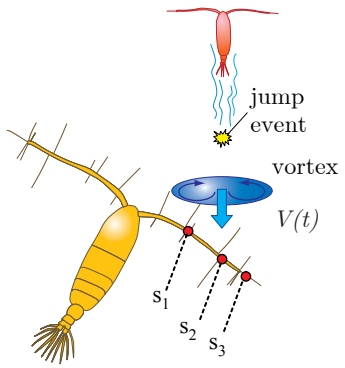


Figure 3.33: A predatory copepod can potentially gain higher order information from the detection of a vortex shed from an escape reaction of, for instance a copepodite. On the planktonic scale, the size and velocity of the vortex changes only slowly over time and distance travelled. Furthermore, high Re flows can maintain sharp velocity gradients. Thus, the measured size and velocity of the vortex relates to the jump speed and size of the escaping copepod. The timing of the signal as it passes different setae, $s1$, $s2$ and $s3$, gives directional information, pointing along the path of the escaping copepodite.

A single jump or a rapid swimming bout is initiated when the copepod applies a thrust on the fluid. While more sophisticated models of how copepods generate thrust have appeared (*e.g.* Morris et al (1985)), for simplicity I assume this is affected by a single impulse: a constant force F_T applied for a finite time interval, Δt. During this time the copepod accelerates, comes under the influence of drag, and approaches 'terminal' velocity, U_0. When the copepod stops applying thrust, it decelerates due to drag, and finally comes to a stop. We can model the thrust F_T by considering the rate of change of momentum of fluid pumped by the copepod's swimming appendages. Specifically, if the velocity of pumped fluid relative to the copepod is u, and is pumped through a cross sectional area A, then in the high Reynolds number regime $F_T = \rho A u^2$ where ρ is the density of the fluid. However, there is a maximum rate U at which the copepod can pump water. This is physically fixed by the rate at which the copepod can move its swimming appendages. Thus the relative velocity pertinent to the thrust is $u = U - v$ where v is the speed of the copepod through the water. The thrust that can be applied has a maximum when the copepod is stationary, and falls off as the copepod accelerates. Rowers are familiar with a similar process in starting a rowing boat. For simplicity, setting the density of the copepod equal to that of the fluid, and $A = \pi a^2$, the equation of motion for the copepod can be written as

$$\frac{dv}{dt} = \frac{3}{4}\frac{(U-v)^2}{a} - \frac{d}{\tau}v \tag{3.6.24}$$

Here $v(t)$ is the copepod's velocity through the water and a (in this case) is the equivalent radius of the cross sectional area of the copepod perpendicular to its direction of motion. $\tau = 2a^2\rho/(9\mu)$ is the Stokes' time scale which for a 100μmparticle is about 1 ms. The drag factor d is

$$d = 1 + 0.1315 Re^{(0.82 - 0.05 log_{10}(Re))}$$

(Clift et al. (1978), reviewed in Fung (1998)), and covers the range of Re up to O(10^3).

If the initial momentum impulse supplied to the fluid is strong enough, a vortex will be shed, propagating aft of the copepod and entraining fluid as it moves. Theoretical considerations (Maxworthy (1972), Shariff & Leonard 1992), suggest that the time evolution of the translational velocity, V, of the vortex is of the form

$$V(t) = A(t + t_0)^{-1}$$

where t_0 is a "virtual" time origin, and A is a constant.

As an example, the data reported by van Duren (2000) on the vortex shed by an escaping *T. longicornis* copedodite stage I (length 330 μm) yields a reasonable fit for $A = 2$ mm and $t_0 = 0.1$ s. The dimension of the vortex can be estimated by noting that the total momentum of the vortex must be conserved. Thus, if b is the radial dimension of the vortex, Vb^3 is a constant. Defining X as the distance of the vortex from the point of the initial impulse, we can write

$$X = A\ln(1 + t/t_0) \tag{3.6.25}$$
$$V = V_0 \exp(-X/A) \tag{3.6.26}$$
$$b = b_0 \exp(-X/(3A)) \tag{3.6.27}$$

where $V_0 = A/t_0$ is the initial velocity and b_0 is the initial radial dimension of the vortex. It is reasonable to assume that b_0 scales with the size of the organism generating the vortex. By equating the initial momentum imparted by the copepod to the fluid with the initial conditions of the vortex, we can approximate $3U^2a\tau \approx 8V_0b_0^3$. Thus, by combining the motion of the copepod and the motion of the vortex we have a potential means of predicting the escape success as a balance between the ability to flee danger (jump length), and the exposure to attack risk (hydromechanical signal of vortex).

A few salient points can be drawn from a scale analysis of the processes involved. The time scale associated with small vortices is long compared to typical time scales for copepod jumps. Even for the relatively weak vortex generated by *T. longicornis* CI as reported by van Duren (2000), velocities in excess of 1 mm/s were evident 1 s after the impulse event. The Stokes time scale for *T. longicornis* CI is of the order 1 ms and the escape interval is of the order 10 ms. In this time interval, the vortex is hardly diminished still having 90% of its initial velocity. Thus a predatory copepod with sensitivity of $s = 100$ m/s can sense the vortex up to 10 mm from its origin, for 20 s after the jump event, and over a volume with radius 2 mm. This begs the question; will not a large powerful eddy confuse the predator rather than alert it to a potential prey? This may very well be the case for a relatively large prey generating a correspondingly large vortex. However, small-scale vortices remain coherent over quite a long distance. Taking *T. longicornis* CI as an example again, then the size of the vortex has increased less than 20% over a travel distance of 1 mm, and doubles in size only after more than 4 mm travel distance. Thus the vortex carries information on the size of the organism that generated it over relatively long distances. Not only that, but the coherence of the vortex suggests that directional information is potentially available to the predator Fig. 3.33. Seen in this way, a single escape jump for *T. longicornis* CI is a high-risk manoeuvre.

3.7 Fluid drift and moving bodies

As a body moves through the water, (*e.g.* a swimming organism, sinking faecal pellet, rising gas bubble), a fraction of the fluid it passes through is dragged along with it, causing a general drift of the fluid in the direction in which the body moves. This process has recently been proposed as a possible means by which marine organisms promote the mixing of the oceans (Katija and Dabiri, 2009).

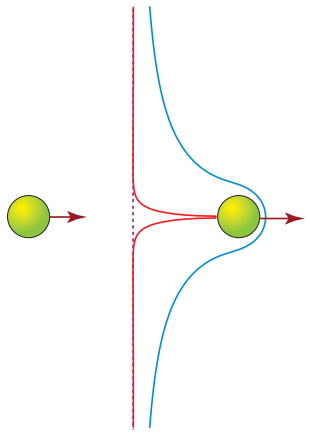


Figure 3.34: The drift associated with a horizontally translating spherical body for the limits of stokes flow ($Re \to 0$ blue line) and potential flow ($Re \to \infty$ with no vorticity, red line). The drift is represented by the horizontal displacement of fluid elements, initially in a vertical line (dashed line) when the sphere is 5 radii to the left. I use a 4th order Runge-Kutta integration scheme to solve in spherical polar coordinates:

$$\dot{r} = \psi_\theta/(r^2 \sin(\theta))$$

$$\dot{\theta} = -\psi_r/(r^2 \sin(\theta))$$

where the stream function ψ is

$$\psi = -\frac{1}{2}U(r^2 - \frac{3a}{2}r + \frac{a^3}{2r})\sin^2(\theta)$$

for stokes flow, and

$$\psi = -\frac{1}{2}U(r^2 - \frac{a^3}{r})\sin^2(\theta)$$

for potential flow.

3.7.1 Fluid drift for a body force

The fluid drift caused by moving bodies is active not only at low Reynolds number where viscous drag is an evident mechanism, but also at high Reynolds number. The high Re affect is perhaps less obvious, but has to do with the added mass of the moving body; the effective mass of fluid that has to be parted to allow the body to move forward. Theoretical results suggest that the drift volume for a Stokes flow in an unbounded fluid is infinite; a result that requires $Re = 0$ which while a reasonable approximation in the vicinity of a moving body, looses credibility at larger distances (Eames et al., 2003). For the case of potential flow (*i.e.* high Re and irrotational) in an unbounded fluid, the drift volume apparently approaches that of the added mass; an estimate sometimes known as Darwin's proposition (Eames et al., 1994) in deference to Darwin (1953)[15] and to the fact that a rigourous proof remains elusive.

To illustrate this effect we can consider the 2 Re extremes for a sphere of radius a moving at velocity U through a fluid in response to a body force (Fig. 3.34). This show the horizontal displacement of an initially vertical line of ideal tracers through which the sphere moves. The drift volume is then indicated by the integral between the final and initial positions of the tracer. It should be noted that the estimate of drift volume depend on where the sphere starts (*i.e.* when the traces are aligned vertically) and the end time of the simulation. Ideally the time span should be $[-\infty, \infty]$, which leaves the mathematical formalism as well as its physical interpretation in a conceptual limbo. In a practical sense though, a simulation over a few body lengths adequately demonstrates the mechanism, with the Stoke's drift clearly greater than the potential drift. The drift illustrated here would be appropriate at low Re for sinking particles and bubbles. But what of swimming organisms? How will the thrust they apply on the fluid effect the drift?

[15] Charles Galton Darwin (1887–1962) grandson of Charles Robert Darwin

3.7.2 Fluid drift for a self-propelled body

A suitable idealized model to address the drift associated with a swimming organism is the Stoke's flow associated with a squirmer. This has the general characteristics of a stresslet reflecting the two opposing forces acting on the fluid, drag and thrust, as well as the appropriate boundary conditions for a translating sphere (Sec. 3.4.4). In Fig. 3.35 are estimates of the drift associated with 2 squirmers, one with $\kappa = 0.5$ and the other with $\kappa = -0.5$ corresponding to a "pusher" and "puller" respectively. Remarkable is the observation that while these are only somewhat different from each other, they are substantially different from the Stokes' flow for a similar sized sphere moving under the influence of a body force. Indeed, these simulations are more like the potential flow (*i.e.* inviscid and irrotational).

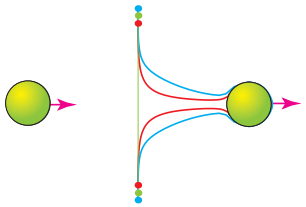


Figure 3.35: The drift associated with a horizontally translating squirmer with $\kappa = -0.5$ (blue line) and $\kappa = 0.5$ (red line). The drift is represented by the horizontal displacement of fluid elements, initially in a vertical line (dashed line) when the sphere is 5 radii to the left. The solutions scheme is as in Fig. 3.34 where the stream function ψ is

$$\psi = -\frac{U}{2r^2}\sin^2(\theta)(a-r)(3a^2(a+r)\kappa\cos(\theta)-(a^2+ar+r^2))$$

Mathematically this is not surprising as the potential flow solution figures in the squirmer solution; the potential flow solution and squirmer solution are identical when $\kappa = 0$.

3.7.3 Fluid drift and stratification

Where fluid drift takes on a potentially important environmental role is in the vertical motion of solid bodies through a stratified fluid. Such vertical motion may include the sinking of detrital material or the vertical migration of zooplankton. You can think of Fig. 3.34 – 3.35 rotated through $\pm 90°$ where the line of tracer particles represents the interface between two layers of contrasting density. Drift thus transports fluid of different densities vertically where it may be subsequently mixed with ambient water so promoting destratification of the oceans (Katija and Dabiri, 2009). While the effect per individual may be small, the concerted migration of schools of zooplankton may have an appreciable effect. However, before such a mechanism can be taken as a credible contributor to biomixing of the oceans, there are some issues that have to be addressed. Most importantly, fluid drift in and of itself is not mixing. While certain volumes of fluid may be transported vertically, this is a transient effect; dense water transported upwards for instance will eventually slough-off the moving body as it comes under the influence of gravity, and fall back, dissipating its potential energy as internal waves. This is a phenomenon that has been observed in the lab, where particles sinking through a density layer will slow, stop or even bounce before continuing downwards; the "attached" low density water momentarily reducing the particles effective settling velocity until it becomes "detached".

4 Turbulence, Stirring and Mixing

Ocean turbulence as it blends and disperses, effects the distribution and abundance of marine plankton. It is of more than passing interest for plankton ecology. Turbulence stirs the oceans – it mixes cold nutrient rich water upwards, fertilizing the sunlit surface ocean, it pulls blooms and solute patches into filaments and tendrils, and it brings about relative motion between individual plankters thus promoting their interactions.

Turbulence is fluid motion unplugged – it is where an unrestrained fluid reveals its true nature. Turbulence, it is said, is one of the last great challenges of physics. This is somewhat of an overstatement. While turbulence may be resistant to a complete and comprehensive physical description, it is not completely enigmatic. Everyone, after all, has some intuitive feeling for what turbulence is. Further, it does yield to predictability at least in a statistical sense, and its fundamental dynamic laws as expressed in the Navier-Stokes equations, are known. It is just very, very complex[1].

This complexity can be rendered manageable, at least for plankton dynamics, by focusing on the turbulent phenomena relevant for specific aspects of the marine ecosystem. That is, the effect of turbulence on a whole phytoplankton bloom may be quite different from than responsible for the vertical exchange of nutrients into the bloom, and may be different again from the turbulent process that govern the contact rates of a zooplankter feeding on members of the bloom community. Turbulence is one process but many phenomena.

Some relevant aspects of turbulence phenomenology can be viewed in terms of a particle-number hierarchy. Specifically, how is a single particle (*i.e.* fluid element, dissolved tracer or neutrally buoyant plankton) moved in space, compared to how a pair of particles move relative to each other, compared to how a cluster of particles changes shape. All of these phenomena may be active in the same turbulent event, yet different biological processes will be subject to each one of these different aspects.

In this chapter I will start with the physical aspects of turbulence, what it is, how it can be characterized, measured and modelled. There follows a discussion on the implication of turbulence on the life experience of individual plankton. How turbulence lends structure to the marine environment and how this shapes the behavioural ecology of zooplankton. For instance, turbulence increases encounter rates on the one hand, but can impair detection ability and capture efficiency on the other. Chemosensory detection is also influenced by turbulence as patches and plumes of solutes are deformed and mixed. I finish with a discussion on biomixing of the oceans – a novel but flawed idea that swimming micro-organisms (zooplankton included) can lend substantially to the mixing of the world's oceans.

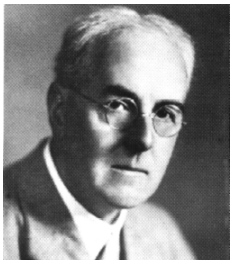
[1]Both Werner Heisenberg and Sir Horace Lamb apparently saw turbulence as one of the two topics to interview God on when they died. They didn't agree on the other topic though, Heisenberg opting for relativity, while Lamb would choose quantum electrodynamics.

4.1 Physics of ocean turbulence

Turbulence in the oceans arises because the forces that cause water to move are generally organized at a much greater spatial scale than the scale at which the flow is subject to friction. What defines large scale and small scale in this context is the relative contribution of inertia (the tendency for a fluid or any body to continue to move) and friction. This concept is captured by the Reynolds' number $Re = UL/\eta$ (already encountered in Sec. 3.1.4) where U is the speed of the motion, L is the scale over which the motion is organized, and η is the kinematic molecular viscosity of the fluid ($\eta = 10^{-6}$ to $2 \times 10^{-6} m^2/s$ for seawater). When Re is much larger than 1, inertia dominates and the flow can be turbulent. Only when Re is close to 1 does friction become important. The oceans for instance, are forced by tides, storms, and differential heating and cooling, at scales of many hundreds of kilometres. The speed of motion at these scales is of the order of 1 m/s (*e.g.* a tidal current, the speed of the gulf stream). Thus, at this scale Re is huge, 10^8 or higher; the ocean at this scale is turbulent, and the effect of friction is negligible. However, the ocean does not continue to build up kinetic energy. The only way for it to be dissipated is for it to become organized in smaller and smaller scales, eventually reaching some small spatial scale where $Re \to 1$ and friction becomes significant.

The concept of the turbulent energy cascade, first proposed by Richardson[2], and formalized by Kolmogrov[3], is truly one of the great advances in physical sciences of the 20$^{\text{th}}$ century. The theory, elegantly simple (*cf.* Sec. 4.1.1) predicts that turbulent energy is distributed with respect to scale ($\ell = 2\pi/k$) as $E(k) \sim k^{-5/3}$, and is controlled by 3 parameters, the integral length scale (the scale at which external mechanical energy is supplied to turbulent motion), the viscosity of the fluid, and the dissipation rate. Furthermore, these characteristics are universal; they govern turbulence in a stirred cup of coffee, the oceans as well as in interstellar gas clouds.

4.1.1 Richardson's cascade and Kolmogorovs scales


Lewis Fry Richardson

In the 1920s and 30s Sir Lewis Fry Richardson came up with the concept of turbulent eddies, and the energy cascade – namely that large scale eddies (*i.e.* coherent inertial structures of the fluid flow) become unstable and break into smaller eddies which break into smaller eddies still. This process of eddy breaking transports kinetic energy from large scales where it is dumped by the mean flow (*e.g.* wind driven currents, tides, waves) to smaller scales where it is eventually dissipated to heat by the fluids viscosity – a concept neatly summed up in the couplet:

> Big whirls have little whirls that feed on their velocity,
> and little whirls have lesser whirls and so on to viscosity.

[2]Lewis Fry Richardson (1881 – 1953) English mathematician, physicist and meteorologist. One of the first to attempt numerical weather forecasting, he was a lifelong pacifist, writing "The Statistics of Deadly Quarrels" published in 1950. Apart from his many contributions to the theory of turbulence, he was also a founder of the field of fractals, becoming interested in the measure of coastlines and borders as a measure of the tendency of nations to go to war.

[3]Andrey Nikolaevich Kolmogorov (1903 – 1987), Russian mathematician who worked extensively in classical mechanics, probability theory and turbulence. His name is associated with a whole range concepts including the Kolmogorov-Arnold-Moser theorem predicting the behaviour of quasi periodic motion in classical mechanics, the Kolmogorov microscales in turbulence, and the Chapman-Kolmogrov master equation in stochastic processes.

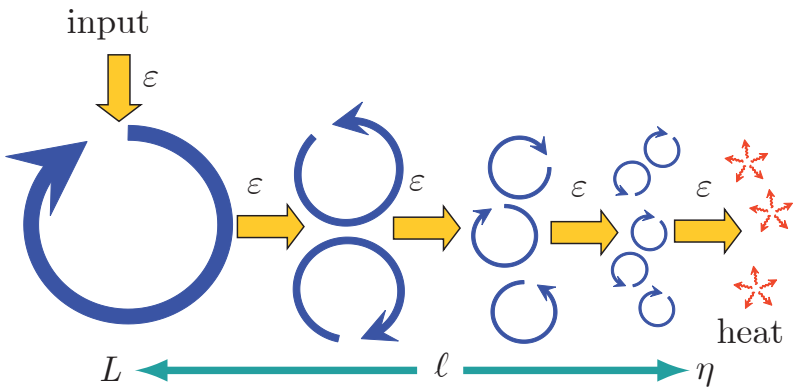


Figure 4.1: The turbulent kinetic energy cascade as envisaged by Richardson and Kolmogorov. Kinetic energy is delivered to the fluid at some large scale L at a rate ε. It cascades to ever smaller scales to be eventually dissipated as heat. This happens at scales of the order of the Kolmogorov scale η.

Kolmogorov (1941) in a very short paper, proposed that the characteristics of the energy cascade are governed solely by ε, the rate at which energy is transferred from scale to scale, and that any imprint of processes acting at larger scales are scrambled, giving turbulence an isotopic and steady aspect. The characteristic time scale of an eddy of size ℓ and kinetic energy u_ℓ^2, is $\tau_\ell \approx \ell/u_\ell$. *Assuming* that this is the time within which the eddy looses its kinetic energy to smaller scales [4], the energy dissipation rate can be written

$$\varepsilon \approx \frac{u_\ell^2}{\tau_\ell} \approx \frac{u_\ell^3}{\ell} \tag{4.1.1}$$

From this it follows that the velocity scale is

$$u_\ell(\ell) \approx (\varepsilon\ell)^{1/3} \tag{4.1.2}$$

Of course, ε is not just the rate energy from eddies of size ℓ is transferred to smaller scales, but, in order for no energy to build up, it is also the rate at which energy cascades across all eddy sizes, from the very largest to the very smallest. The very smallest eddy size, $\ell = \zeta$, is set where viscosity comes into play; that is where the Reynolds number $Re = u_\ell(\zeta)\zeta/\eta = 1$. Inserting into Eq. 4.1.2 and rearranging gives

$$\zeta = \left(\frac{\eta^3}{\varepsilon}\right)^{1/4}. \tag{4.1.3}$$

This is termed the Kolmogorov length scale, and is generally taken as the lower limit of the inertial sub-range. Other related scales are the Kolmogorov time scale $\tau_K = (\varepsilon\zeta)^{1/3}$ and velocity scale $u_K = (\zeta^2/\varepsilon)^{1/3}$.

The energy spectrum (*i.e.* the turbulent kinetic energy density per unit wave-number) then follows: $u_\ell^2 \approx (\varepsilon\ell)^{2/3}$ and wave number $k = 2\pi\ell^{-1}$, so that

$$E(k) \approx \ell u_\ell^2 = \alpha_K \varepsilon^{2/3} k^{-5/3} \tag{4.1.4}$$

[4] The reasoning here is that size size and kinetic energy are governing variables, there is only 1 combination that gives the dimensions of time. So to within a constant, the characteristic time scale for energy transfer ℓ/u_ℓ

where α_K is a constant (often given the value 1.5). This is the famous $-5/3$ energy spectrum law.

While the formulation of Kolmogorov's spectrum is often criticized as *ad hoc*, (it is based largely on dimensional analysis, *cf.* Fig. 4.2 for summary), its general characteristics have been observed over a wide range of natural turbulent settings. The first comprehensive

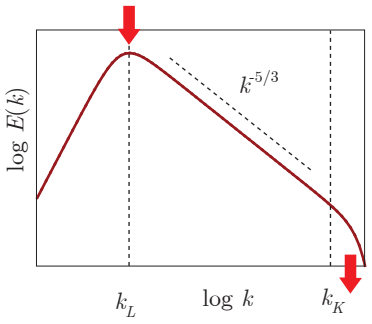


Figure 4.2: The turbulent energy spectrum as a function of wave number k. Energy is imparted at large scales (small k and dissipates at small scales (large k). Between these limits – the so called inertial scale, energy cascades at a rate, according to Kolmogorv's theorem, determined solely by the energy dissipation rate ε. Based on dimensional analysis, the dependence of the energy content per unit mass in a small wave-number band ($E(k)$ is the energy per unit mass (units L^2T^{-2})) between k and $k + dk$. Since k has units $k \sim L^{-1}$, $E(k)$ has units, $E \sim= L^3T^{-2}$ and $\varepsilon \sim L^2T^{-3}$ the only way of combining E, k and ε that maintains the correct dimensionality is $E \sim \varepsilon^{2/3}k^{-5/3}$.

confirmation of the shape of the turbulent spectrum came for measurements in Knight Inlet, a tidal channel off the west coast of British Columbia (Grant et al 1962), and has since been documented in many other parts of the oceans (Gargett et al., 1984; Gargett, 1989; Yamazaki et al., 2002), atmosphere, as well as laboratory studies.

4.1.2 Richardsons law

A direct consequence of the turbulent energy spectrum is a relationship describing how ideal traces in a turbulent flow will move relative to each other in a statistical sense (Fig. 4.3). This is Richardson's law, and is perhaps the most useful of all phenomenological descriptions of turbulence for marine ecology.

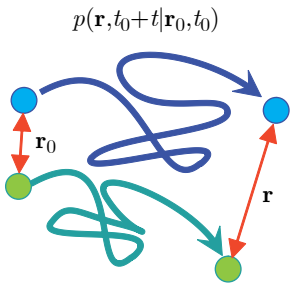


Figure 4.3: Richardson's law is based on the 2 particle phenomenology of turbulence. It formalizes the probability that two particles are separated by a distance r at a given time $t_0 + t$, given that they were separated r_0 at time t_0. The principle can be extrapolated to many particles at least at short time scales, where one particle is the centre of mass of the cluster, and the other is a test particle within the cluster. The ensemble average of many such test particles gives an expression for the rate of change of variance of the cluster within a turbulent flow.

A patch of solute or a puff of gas in a turbulent flow will deform, disperse, and eventually dissipate. The rate at which such a patch disperses is of central interest in many applied problems in turbulence. Consider a cluster or neutrally buoyant particles (ideal tracers or fluid elements) in a turbulent flow. We may for instance define the centre of mass of the cluster $\mathbf{x}_0 = \langle \mathbf{x}_i \rangle$ and examine how the variance of the particle displacement $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_0$ changes in time. The cluster thus has a characteristic scale $r = (\langle \mathbf{r}_i \cdot \mathbf{r}_i \rangle)^{1/2}$. Given the eddy structure of turbulence, we can immediately reason that eddies much larger than the patch

($\ell \gg r$) will not contribute to the *relative* motion. These eddies will simply move the cluster as a whole. Much smaller eddies ($\ell \ll r$) will deform the cluster, but the kinetic energy contained in these is relatively low. Thus, it is those eddies that are commensurate with the cluster size ($\ell \approx r$) that affect the greatest degree of dispersion. We have already seen that the velocity scale associated with eddies of scale ℓ is $u_\ell \approx (\varepsilon \ell)^{1/3}$ (Sec. 4.1.1). That is, if we take r as the size of a patch, then the rate at which r increases is

$$\frac{dr}{dt} = \alpha_R (\varepsilon r)^{1/3} \tag{4.1.5}$$

where α_R, Richardson's constant, turns out to be close to unity[5] (Lesieur, 1997). Similarly, the variance of the displacement can be estimated to vary as $r^{4/3}$. This is expressed in Richardson's law;

$$K(r) = \frac{1}{2}\frac{dr^2}{dt} = r\frac{dr}{dt} = \alpha_R \varepsilon^{1/3} r^{4/3} \tag{4.1.6}$$

where $K(r)$ is the rate of change of variance of a patch of size r lying in the inertial sub range $[L > r > \eta]$. That is, the rate of change of variance of a cloud of marked fluid elements – its diffusivity (*cf.* Sec. 2.2.2) – increases faster the larger it gets, to the 4/3 power. Simply put,

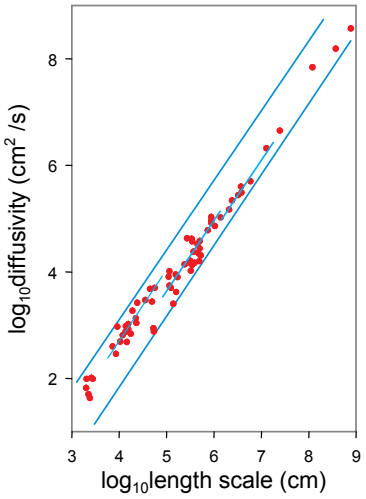


Figure 4.4: Empirical evidence of Richardson's law, diffusivity $K(r)$ versus length scale r. The observations come from Okubo (Okubo, 1971)where the size of a dye patch, released under various conditions, was measured over time. The envelop of results conform to a $r^{4/3}$ power law, although all observations taken together have a somewhat flatter relationship. A reinterpretation of these observations by Ozmidov's generalized power spectrum (Ozmidov, 1965) – where it is recognized that turbulent kinetic energy can be input at multiple scales (*i.e.* multiple L's) – provides a piecewise $r^{4/3}$ power law fit (Okubo and Ozmidov, 1970).

when particles are close together, they are pretty much synchronized to the same large scale eddies and relative motion is due only to eddies similar in size to their separation distance. As they move apart, larger eddies become involved in their relative motion contributing to the rate at which variance increases.

Lastly, it is possible to solve Eq. 4.1.6 in terms of time to give

$$r^2 = \left(\frac{2}{3}\alpha_R\right)^3 \varepsilon t^3 \tag{4.1.7}$$

That is, in a turbulent flow, variance increases at time to the 3rd power, in contrast to Brownian motion (*i.e.* scale independent diffusion) where variance increases linearly with time. Any process wherein variance is a non-linear function of time is termed "anomalous diffusion", and processes where this increase is faster than linear as in the case of turbulence is termed "super diffusion".

[5]Precise numerical values of Richardsons constant α_R vary between 0.7 and 2.1 although a value of 1.8 is most often applied.

4.1.3 Intermittency and coherent structures

Much of the phenomenology of turbulence is based on some sort of averaging. The turbulent energy spectrum is, for instance, the energy spectrum in some volume of fluid, averaged over many realizations of the same situation. In general, whatever the "averaging" is, it is never equivalent to a proper ensemble average (over all possible states of the system). Ergodicity is usually invoked to replace it by time and/or spatial averaging. These tools are sufficient to reveal some of the most important "universal" features of turbulence.

However, no single realization of a turbulent flow ever looks like its ensemble average. The energy cascade, for instance, is not smoothly continuous with a 5/3 slope across all scales at all times. Energy briefly builds up at one scale to be shed as a pulse running down the cascade. That is, energy transport is intermittent. The spatial corollary to this is that in any realization, turbulence is not evenly distributed through out the fluid, but comes in localized zones – coherent structures – that evolve and dissipate. Billows and swirls are evident everywhere we look in turbulence, and were the particular features that attracted Leonardo Da Vinci's eye. This in itself would not pose a great challenge for the phenomenology of turbulence where it not for the fact that these events – intermittency in time and coherent structures in space – are not normally distributed. This non-Gaussian distribution of events appears to be a common characteristic of many complex systems. Like tornadoes and maelstroms, infrequent but intense events can radically change an average. More challenging to Kolmogorov's theory is that intermittency calls into question its universality.

As a practical matter, intermittency and coherent structures can be largely ignored in the first instance. The relatively robust empirical evidence on a large class of turbulent phenomena supporting Kolmogorov's theory and the predictions made by it stand witness to its power. This is not to say these matters should be ignored entirely. There are instances, particulary in the small scale process of intimate importance to individual plankton, where coherent structures may play a role. These can be treated as stand alone effects without a comprehensive re-analysis of Kolmogov's theory.

4.1.4 Turbulent dispersion in the ocean

If we were to track the distance ℓ between 2 ideal tracers (vanishingly small fluid elements or dissolved molecules) in a turbulent ocean we would find that on average, their separation distance would increase with time. The rate of increase however would depend on their instantaneous separation distance. That is, the turbulent dispersion of tracers in a turbulent flow is scale dependent.

In general, dispersion falls into 4 regimes (Fig. 4.5) depending on scale. These are delineated by 3 intrinsic length scales: L the integral length scale – the length scale of the energy containing turbulent eddies, ζ the Kolmogorov scale – where viscous effects come into play (*cf.* Eq. 4.1.3), and β, the Batchelor scale – the scale at which molecular diffusion comes into effect – given by

$$\beta = \left(\frac{D^2\eta}{\varepsilon}\right)^{1/4} = \zeta\sqrt{\frac{D}{\eta}} \tag{4.1.8}$$

At very small ($\ell < \beta$) and at very large ($\ell > L$) separation distances, dispersion will have a Brownian characteristic in that the increase in variance ($\langle \ell^2 \rangle$) is linear in time. The pre-factor in this increase, however, will be quite different for these 2 regimes, with molecular diffusivity

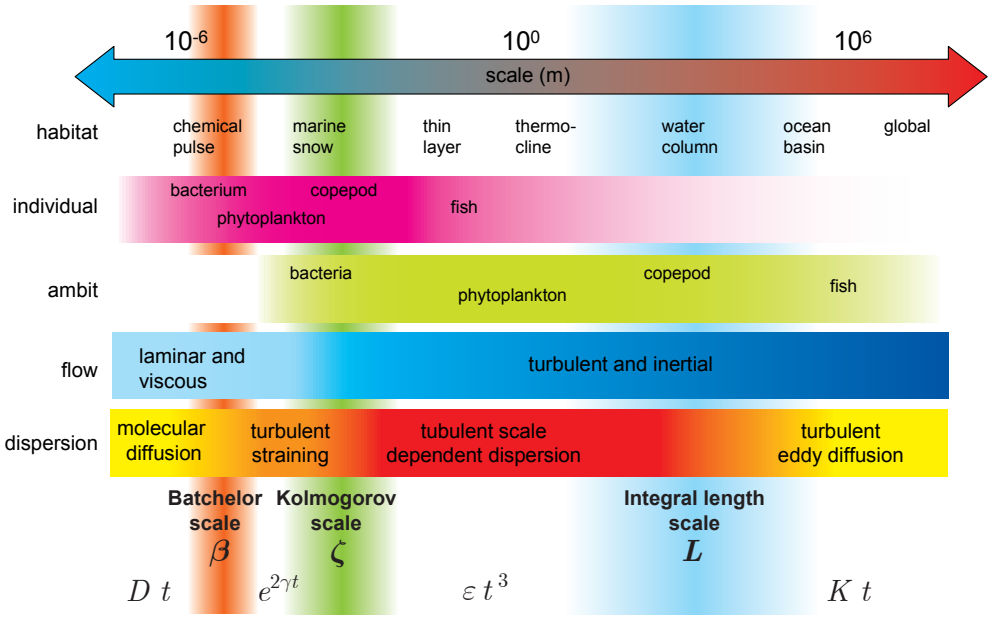


Figure 4.5: In the marine environment, physically and biologically relevant length scales range over 12 decades from microns (individual bacteria) to thousands of kilometers (ocean basins). These length scales are set by the size of the individuals, the habitats they inhabit, and the scales a typical individual within a population may cover over its life time. Different aspects of turbulence influence different aspects of the interactions between individuals, and within communities and populations. Turbulent dispersion – the relative motion of chemical tracers and passive planktonic particles – fall into 4 regimes, where the rate of change of variance (*i.e.* diffusivity) scales differently with time and spatial scale.

(D) controlling small scales, and turbulent eddy diffusivity (K) at large scales. Between these 2 extremes, are 2 additional regimes. Within the inertial sub-range ($\zeta < \ell < L$), variance increases with t^3 according to Richardson's law. Finally, there is a small range between the Kolmogorov and Batchelor scales ($\beta < \ell < \zeta$) where the separation is controlled by shear, and variance increases exponentially according to $\exp(2\gamma t)$ where γ is the shear rate.

Turbulent dispersion and diffusion: Mixing within a turbulent flow leads to the dispersion and diffusion of suspended material and water characteristics. Turbulent dispersion and turbulent diffusion are pretty much the same thing. In an applied sense, dispersion tends to be reserved for what happens to particles, while diffusion for what happens to molecules and heat. This has lead, in some peoples minds, to the notion that they are distinct, separate aspects of turbulence. They are not. Turbulent diffusion is equivalent to the turbulent dispersion of an ideal tracer with vanishingly small molecular diffusivity. Under the fairly reasonable assumption that turbulent dispersion and molecular diffusion are statistically independent, this can be stated even more succinctly: Turbulent diffusion is equivalent to the turbulent dispersion of an ideal tracer plus molecular diffusion.

Statistical Theory of Diffusion: It is almost intuitive that random motion, whether due to turbulence or molecular collisions, leads to a dispersion of material. How this dispersion is related to the underlying kinematics of movement is the subject to the statistical theory of diffusion. This theory was built up from statistical mechanics (Maxwell, Einstein, Taylor) in concepts already encounter in discussing Brownian motion.

Turbulent eddy diffusivity: A natural extension of Richardson's law (Eq. 4.1.6), and the definition of the integral length scale L lead to

$$K = \alpha_R \varepsilon^{1/3} L^{4/3} \tag{4.1.9}$$

4.2 Measuring turbulence

Given that turbulent motion is prevalent across a multitude of scales, it may seem a daunting task to either (i) establish a meaningful parameter describing the turbulent state, and (ii) devise some means of measuring this parameter. Part of the answer, and that which is most often applied, is to estimate the turbulent dissipation rate from small scale velocity fluctuations.

The average rate at which energy is dissipated – irrespective of details of the flow – is given by

$$\varepsilon = \frac{1}{2}\eta\langle e_{ij} e_{ij}\rangle \tag{4.2.1}$$

where e_{ij} is the rate of strain tensor (*cf.* Sec. 3.1.2) at some point in the flow, and $\langle \ldots \rangle$ represents the ensemble average of the enclosed quantity (*cf.* Sec. 2.2.2). For 3 dimensional isotropic turbulence, the symmetries in e_{ij} lead to

$$\varepsilon = \frac{15}{2}\eta \left\langle \left(\frac{\partial u}{\partial z}\right)^2 \right\rangle \tag{4.2.2}$$

where u is a velocity component orthogonal to the direction z. That is Eq. 4.2.2 expresses turbulent dissipation rate – and therefor the kinetic energy cascade rate across the entire inertial subrange – in terms of one component of the micro-scale shear.

Measurement of the micro-scale shear can be carried out in principal by applying Taylor's frozen turbulence hypothesis. That is, an instrument moving through a turbulent flow fast enough (at speed U say) will sample velocity structures within the flow faster than they evolve themselves. Thus time and distance can be interchanged ($z = Ut$). The same is true if the instrument is fixed and the flow has a mean component U considerably greater than the rms turbulent velocity. Thus small scale shear can be derived from high resolution temporal measurements.

Over the years, 3 main types of technologies have been applied to this problem. The earliest of these were "hot wire anemometry" where the resistance of a heated wire varies with the rate at which it is cooled by fluid flowing over it. This cooling is a function of the speed at which the fluid flows, and the response of the instrument is rapid enough to measure turbulent fluctuations in the flow. This was superseded by piezo-electric shear probes – where velocity fluctuations apply mechanical stress on the sensor giving rise to an electric potential (the same principle as an old-fashioned gramophone needle). Both the fact that these sensors can be tailored to different components of the velocity field, and their high sensitivity means they are the most common means of measuring turbulence in the ocean. (Observations of dissipation rate presented in Fig. 4.12 were taken using a free falling shear probe of this type). Running parallel has been the development of acoustic techniques that rely on the Doppler shift of emitted and reflected sound from suspended particles to give a measure of along beam fluid speed.

In principal these techniques would give estimates of the small scale shear; shear at scales less than the Kolmogorov scale, *i.e.* below 1 cm or so. In practice all of these techniques fall somewhat short of required resolution. In practice, the usual technique is to appeal to Kolmogorov's theory which for isotropic stationary turbulence leads to a universal shape – the Nasmyth universal spectrum – which depends only on dissipation rate and molecular viscosity. Thus a "fit" of observed turbulent energy contained over a range of scales can likewise provide an estimate of the turbulent dissipation rate.

4.3 Modelling turbulence

In nearly all cases, turbulence models are designed to simulate only one particular aspect of turbulence – an all encompassing model is simply beyond our means. Those model that come closest to simulating actual turbulence are Direct Numerical Simulations (DNS). However, the computational cost of these is enormous, going up with the cube of the Reynolds number. Today, DNS with Re up to about 10^4 have been attained, although most oceanographic flows have $Re \sim 10^9$ or greater.

The most common framework for ocean modelling is the Reynold's Averaged Navier-Stokes' equations (RANS) where the effects of turbulence on the flow is encapsulated in the eddy viscosity. Precisely how the eddy viscosity is estimated is the so-called Turbulence Closure problem. The simplest closure is to assume a constant eddy viscosity although this is far from ideal as, whatever else it is; eddy viscosity is an inherent property of the flow. More successful and widely used these days are closure schemes that rely on quasi-empirical dynamic descriptions of turbulent kinetic energy and dissipation rate ($k-\varepsilon$) or turbulent kinetic energy and mixing length scale ($k-\ell$). (k in this context stands for turbulent kinetic energy and should not be confused with wave-number. I would give it a different symbol were it not for the fact that most references are in terms of $k-\varepsilon$ or $k-\ell$).

Somewhere between DNS and RANS are the Large Eddy Simulations (LES). These are based on the the principle that while the dynamics of large eddies are determined by forcing and geometry, small eddies are self-similar and obey some universal law. LES model large eddies directly, and use a "sub-grid" model for smaller eddies. Like DNS, these are usually formulated in Fourier space, although unlike DNS, they rely on a turbulence closure scheme since the smallest wavelength they consider is larger than the Kolmogorov scale. RANS and LES coupled with turbulence closure schemes form the backbone of simulating real world geophysical fluid dynamics.

Given the complexity of turbulence modelling, it is often useful to engage extremely simple models that simulate fundamental characteristics of turbulence – stirring for instance – without embarking on full dynamic simulations. Such models can be extremely useful in highlighting specific phenomena provided that their limitations are identified.

4.3.1 Kinematic simulations

From a statistical point of view, we know that the fluid of motion small scale turbulence has a specific kinetic energy spectrum but is otherwise randomly directed in 3-D space. A relatively simple means of simulating turbulence is to construct a velocity field using random Fourier modes chosen to conform to given physical constraints (Kraichnan 1970, Fung et al. (1992)). The simulated field is constructed such that its kinetic energy spectrum reflects the turbulent energy spectrum (*e.g.* the classic $k^{-5/3}$ cascade) and that its flow field is non-divergent (as is appropriate for an incompressible fluid). Simulations are specified through

the 3 parameters mentioned earlier: ε, μ and L. Fluid elements or parcels can be tracked by integrating the velocity field. Such models have been used primarily to investigate physical phenomena (Fung, 1993, 1998).

Consider an energy spectrum $E(k)$ describing the distribution of kinetic energy in terms of wave number k. In the Fourier approach, the spectrum is divided into N bands, where the width and representative wave number of the n^{th} band are δk_n and k_n. The turbulent velocity vector field $\mathbf{u}(\mathbf{x}, t)$ at any location $\mathbf{x}$ and time t is described as the sum of Fourier components

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t) &= \sum_{n=1}^{N} \left(\mathbf{a}_n \times \hat{\mathbf{k}}_n \right) \cos(\mathbf{k}_n \cdot \mathbf{x} + \omega_n, t) \\ &+ \sum_{n=1}^{N} \left(\mathbf{b}_n \times \hat{\mathbf{k}}_n \right) \sin(\mathbf{k}_n \cdot \mathbf{x} + \omega_n, t) \end{aligned} \tag{4.3.1}$$

where $\hat{\mathbf{k}}_n = \mathbf{k}_n / k_n$ is the unit vector defining the direction of the n^{th} wave vector $\mathbf{k}_n$, $k_n = |\mathbf{k}_n|$, ω_n is its angular frequency, $\mathbf{a}_n$ and $\mathbf{b}_n$ are the associated amplitude vectors, and the cross term ensures that the flow field is non-divergent. The amplitudes of k_n, $\mathbf{a}_n$ and $\mathbf{b}_n$ are calculated as given below, but their directions are chosen to be random in 3-dimensions.

A simple Kolmogorov energy spectrum

$$E(k) = E_0 k^{-5/3} \tag{4.3.2}$$

is perhaps the most direct approach, although other spectral shapes can also be used (Fung and Vassilicos, 1998; Malik and Vassilicos, 1999; Lewis and Pedley, 2001)). N discrete wave numbers are chosen to lie between $k_1 = 2\pi/L$ and $k_N = 2\pi/\zeta$ where $\zeta = (\eta^3/\varepsilon)^{1/4}$ is the Kolmogorov length scale. To avoid cyclic effects, the discrete wave numbers are distributed in a geometric series from k_1 to k_N.

$$k_n = \frac{2\pi}{\zeta} \left(\frac{\zeta}{L} \right)^{(n-1)/(N-1)}$$

The values of $\delta k_n = (k_{n+1} - k_{n-1})/2$ for $n = 2, \ldots, N-1$, $\delta k_1 = (k_2 - k_1)/2$ and $k_N = (k_N - k_{N-1})/2$. The frequency of the n^{th} mode is

$$\omega_n = 0.4\sqrt{k_n^3 E(k_n)}$$

where the factor 0.4 – the unsteadiness parameter (Malik and Vassilicos, 1999). The amplitudes of $\mathbf{a}_n$ and $\mathbf{b}_n$ are given by

$$|\mathbf{a}_n|^2 = |\mathbf{b}_n|^2 = 2E(k_n)\delta k_n$$

Kinematic simulations compare remarkably well (Fung and Vassilicos, 1998) with more accurate but computationally more demanding direct numerical simulations (Yamazaki et al., 1991). They faithfully reproduce turbulent phenomena such as the turbulent structure function (Fig. 4.6). Given their practicality, kinematic simulations via random Fourier modes have been used to study a number of problems relevant in marine ecology, such as in a re-examination of the turbulent encounter process (Lewis and Pedley, 2000; Mariani et al., 2007), and the dispersion of chemical plumes trailing sinking particles (Visser and Jackson, 2004).

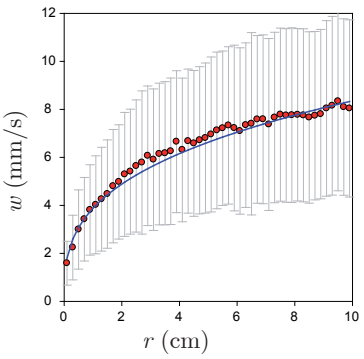


Figure 4.6: The turbulent structure function $w(r)$ at a dissipation rate of $\varepsilon = 10^{-6}$ m^3 s^2)computed by a kinematic simulation compared to the relationship predicted by Richardson's law. In this case Richardson's constant is set at $\alpha_R = 1.8$, which is consistent with empirical measures. The ability of kinematic simulations in faithfully reproducing such phenomena give credence to their use in simulating turbulent encounter rate processes.

4.3.2 The renovating wave

A relatively simple model that can be used to simulate important 2-D aspects of turbulence – in particular stirring and stretching – is the renovating wave model (Young et al., 2001). Most simply put, the "wave" part of the model introduces displacements of fluid elements while the "renovating" part folds them.

In 2D, this motion can be written in terms of a stream function

$$\psi_n(x, y, t) = (U/k)\cos(xk\cos\theta_n + yk\sin\theta_n + \varphi_n) \tag{4.3.3}$$

where U is the velocity scale,k is the wave number, and θ_n and φ_n are randomly varying angles, uniformly distributed on $[\pi, -\pi]$ (more on this later). The stream function renovates at regular time steps in that new random angles θ_n and φ_n are chosen. Integrating over a single time step τ gives an iterative map of fluid particle locations as:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \tau U \sin(c_n x_n + s_n y_n + \varphi_n) \begin{pmatrix} s_n \\ -c_n \end{pmatrix} \tag{4.3.4}$$

where for brevity $c_n = k\cos\theta_n$ and $s_n = k\sin\theta_n$. Particles following this motion diffuse in a Brownian manner in that their variance increases linearly with time: $\langle (x_n - x_0)^2 \rangle \sim n$. While this is not turbulent dispersion as might be expected within the inertial sub-range, it is none-the-less a useful model for turbulent dispersion at scales larger than the integral length scale as may be active, for instance, on plankton blooms in the open ocean (Fig. 4.7).

4.3.3 Turbulence closure

Modelling turbulent processes in shallow regions such as continental shelves, estuaries and coastal areas most often revolves around the question as to how turbulence effects the vertical exchange of dissolved material and momentum. It is, after all vertical rather than horizontal exchange that has the greatest impact on both the physical (*e.g.* water column stratification) and biological (*e.g.* nutrients) aspects of the marine environment. Unlike the previous models, the phenomenology involved here is single particle statistics. That is, the vertical exchange is an Eulerian process that is largely captured in the evolution of the probability distribution function of a tracer and its associated inert and dynamic properties in the fixed reference frame of a turbulent water column.

The basic formulation of the problem follows Reynolds' argument – that irrespective of the precise details of turbulence, its effect is to exchange momentum and material is a manner

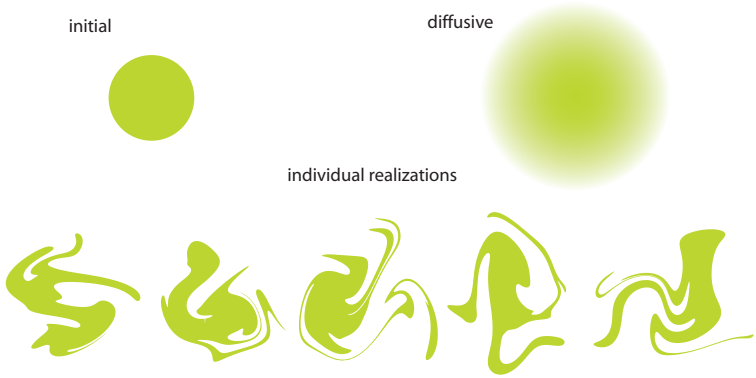


Figure 4.7: To illustrate the renovating wave model, consider the evolution of an initially circular patch. I show here 5 different realizations. This also illustrates an important point. That while in a statistical sense the effect of turbulence is diffusive, no single realization of turbulent stirring has a diffusive characteristic. The concentration of some passive tracer in a small fluid element within a turbulently stirred fluid does not decrease over time as may be expected from diffusion. Rather it remains relatively constant until such time as local variations in the concentration field approach the Bachelor scale.

precisely analogous to diffusion. The coefficients however are strongly influence by turbulence, and are termed turbulent eddy viscosity for the exchange of momentum and turbulent eddy diffusivity for the exchange of material. Loosely defined, the turbulence closure problem is to provide realistic estimates of the eddy viscosity and eddy diffusivity.

The basic formulation can be achieved by thinking of the velocity field being composed of a mean component and a fluctuating turbulent component. Specifically

$$u_i = \overline{u_i} + u_i' \tag{4.3.5}$$

where the overbar represents an averaging operation such that $\overline{u_i}$ is the slowly varying component and u_i' the more rapid turbulent fluctuations. By definition $\overline{u_i'} = 0$. Using the same operator on the Navier-Stokes' equations yields

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} + \overline{u_j'\frac{\partial u_i'}{\partial x_j}} = -\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x_i} + \frac{\mu}{\rho}\frac{\partial^2 \overline{u_i}}{\partial {x_j}^2} \tag{4.3.6}$$

These are the Reynolds-averaged Navier-Stokes' equations for the slow time scale, and they contain an extra term (3^{rd} on the left hand side) due to the nonlinear nature of the advective contribution to acceleration. For an incompressible fluid Eq. 4.3.6 can be rewritten as

$$\rho\left(\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(-\overline{p}\delta_{ij} + 2\mu\overline{e_{ij}} - \rho\overline{u_i'u_j'}\right) \tag{4.3.7}$$

where $\overline{e_{ij}}$ is the averaged rate of strain tensor (*cf.* Sec. 3.1.2)[6]. The quantities $\rho\overline{u_i'u_j'}$ are termed the Reynolds stresses, and they play the same functional role in exchanging momentum as the Reynolds averaged stresses $2\mu\overline{e_{ij}}$. The principal concept in turbulence closure schemes is

[6]This illustrates the principle, but real applications to oceanic conditions generally have to deal with stratified turbulence. That is, there is an additional dynamic term in the pressure gradient due to variations in density: density that itself is mixed by turbulence.

to assume that that the Reynolds stresses are functions of the mean components of the flow, consistent with a relationship of the form

$$\rho\overline{u_i' u_j'} = -2\rho\eta_t(\overline{u_i}, \overline{p}, \mu)\overline{e_{ij}} \tag{4.3.8}$$

The parameter η_t is the turbulent eddy viscosity and will in general depend on properties of both the fluid and the flow – a concept first introduced by Boussinesq in 1877. A similar Reynolds decomposition of the advection-diffusion equations lead to a specific formulation of the turbulent eddy diffusivity.

This is the background. Now the closure itself. While there are many specific closure schemes, I will briefly discuss the $k - \varepsilon$.

If, as assumed above, eddy viscosity is solely dependent on the state of the turbulent flow, then there is really only a few combinations of 2 bulk turbulent variables that can be used to achieve the required dimensions[7] of viscosity *i.e.* $[L^2/T]$. One near universal choice is k, the turbulent kinetic energy. It is a relatively simple matter to write a equation describing the dynamic of turbulent kinetic energy by taking the dot product of the Navier-Stokes' equations and the averaging. The units of k are $[L^2/T^2]$. The other parameter has been variously chosen as the turbulent length scale ℓ or a turbulent frequency ϖ although the parameter most commonly used is the turbulent dissipation rate ε. Unlike length and time scales, turbulent dissipation rate has a clear physical definition, and is measurable. Furthermore, it is possible to write an equation describing the dynamics of dissipation rate (while formally this works, I always found it hard to give a physical meaning to some of the terms – what does the diffusion of dissipation rate mean for instance). The dimensions of turbulent dissipation rate is $[L^2/T^3]$. Dimensional analysis (the Kolmogorov-Prandtl relationship) thus suggests that

$$\eta_t = c_\mu \frac{k^2}{\varepsilon} \tag{4.3.9}$$

where c_μ is a supposedly universal constant ($c_\mu = 0.09$ according to empirical observations). A handful of other "universal" constants in the dynamic equations and for eddy diffusivity can be deduced from dimensional analysis on ensemble averaged processes and evaluated through empirical observations.

Turbulence closure schemes such as $k-\varepsilon$ are the backbone of nearly all modern ocean models. Then have been tested and calibrated over the years and can be thought reliable in most applications.

Random walks and flights: Turbulence closure schemes along these lines provide estimates of turbulent eddy diffusivity – an Eulerian measure. In order to incorporate this into individual-based simulations, tracing particles through turbulent flow, random walks or random flight can be implemented. Such models can be used, for instance, to calculate the exposure of phytoplankton cells to light as they are advected vertically by turbulence. Two points to note however:

- Such implementations however, only address single particle phenomenology and do not describe the *relative* motion of organisms. Turbulence closure models do however provide estimates of turbulent dissipation rate – a measure that can be used in sub-models of two particle phenomenology through Richardson's law.

[7] According to the Buckingham π theorem, a physically meaningful relationship of n physical variables involving m fundamental units (*e.g.* time, space, mass) can be expressed in terms of $n - m$ dimensionless variables.

- Turbulent diffusivity is in general and function of space. As such, naïve random walk and random flight models are inappropriate (Visser, 1997), as will be discussed further in chapter 5

4.4 Turbulence and productivity

For most marine systems, productivity (*i.e.* the rate at which organic carbon is fixed by phytoplankton and transferred to zooplankton or fish larvae populations) has a dome-shaped relationship with turbulence – a little turbulence stimulates production, but too much has a negative effect (Fig. 4.8). For phytoplankton (primary production), a little turbulence can

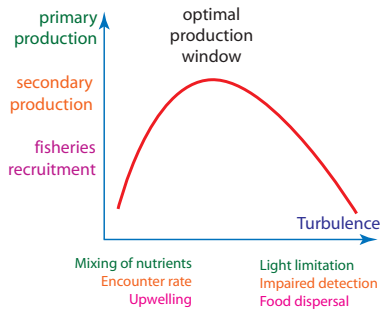


Figure 4.8: The dome-shaped response of primary, secondary and fisheries production in marine systems to turbulence.

mix nutrients into the well-lit surface layer stimulating production, but too much, and the plankton cells themselves become mixed out of the euphotic zone limiting exposure to light so that production drops. Higher up the food chain, increasing turbulence increases the rates at which zooplankton encounter their prey, but high levels impair the ability of zooplankton to detect and capture prey. A similar process is in effect for fisheries recruitment. On an individual predator-prey basis, fish larvae are susceptible to both the positive effects of turbulent encounter rate and the negative effects of impaired capture ability (MacKenzie et al., 1994; MacKenzie and Kiørboe, 2000). Furthermore, on a population level, this can be seen as a response to the dome-shaped relationship of primary and secondary production (Cury & Roy 1989), with stimulated primary production giving a positive effect at low turbulence levels, and food dispersal giving a negative effect as turbulence increases (MacKenzie and Kiørboe, 2000).

4.5 Turbulent encounter rates

Particles, jostled around in a turbulent flow, have an increased chance of bumping into each other – the more turbulent the flow, the more jostling and the higher the chance of encounter. The relevant turbulent measure that determines this encounter rate is the change of the relative velocity w along the line that separates the two particles. As seen in the particle pair phenomenology, this velocity variation is determined by Richardsons law in the inertial subrange. That is,

$$w = \frac{dr}{dt} = \alpha_R(\varepsilon r)^{1/3} \tag{4.5.1}$$

cf. Eq. 4.1.2. The turbulent velocity w is scale dependent. It is relatively low for particles that are close together, but increases at greater distances - a consequence of the relative motion becoming subject to larger and larger eddies.

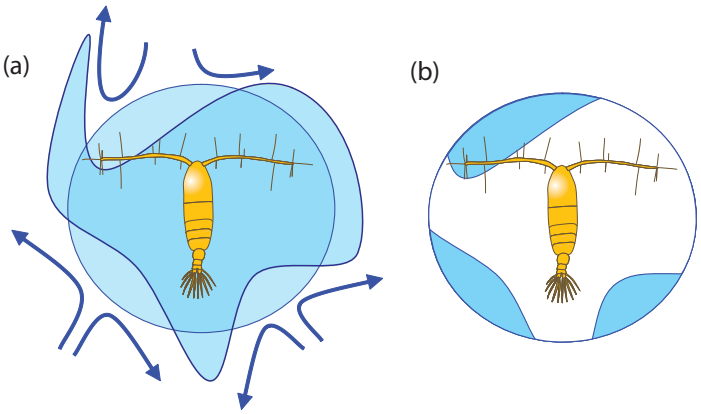


Figure 4.9: Turbulent encounter rate is due to the turbulent deformation of the fluid, carrying "new" water and its associated prey organisms into the predator's detection volume.

How turbulence impacts encounter rates can be most easily seen by considering the turbulent deformation of a spherical blob of fluid (Fig. 4.9). This may, for instance represent the fluid contained within the perception zone of ambush predator. A small time step later, this fluid volume has deformed. The volume of the deformed blob remains the same, and the perception zone of the predator remains spherical. That is, a certain volume of "old" fluid has been swept out of the perception volume to be replaced with equal volume of "new" fluid.

If $\mathbf{w}$ is the turbulent velocity of all fluid elements on the surface of the perception volume, then the total flux of fluid out of perception volume is zero, that is

$$\oint_s \mathbf{w} \cdot \hat{\mathbf{n}} ds = 0 \tag{4.5.2}$$

However, of interest to the encounter process is the rate at which "new" fluid is available for searching.

$$F_{\text{in}} = \frac{1}{2} \oint_s |\mathbf{w} \cdot \hat{\mathbf{n}}| ds = 2\pi R^2 w(R) \tag{4.5.3}$$

The "new" water carries with it potential prey (or mates or predators) that will be encounterer. If C is the concentration of the prey, then the encounter rate is

$$Z = 2\pi R^2 C w(R) = 2\pi R^2 C \alpha_R (\varepsilon R)^{1/3} \tag{4.5.4}$$

This formula is slightly different but of the same form as that derived by (Rothschild and Osborn, 1988), although in this case, the reasoning is hopefully a bit more transparent.

If either the organism or its encounter partner (prey, mate or predator) swim, then encounter rates will increase. The simplest expression including swimming and turbulence is due to (Evans, 1989) and is written

$$Z = \pi R^2 (u^2 + v^2 + 2(w(R))^2)^{1/2} \tag{4.5.5}$$

where u is the swimming speed of the organism, and v the swimming speed of its prey[8]. The relevant turbulent velocity scale $w(R)$ can be conveniently estimated from Richardson's law

[8]Note that Eq. 4.5.4 and Eq. 4.5.5 (with $u = v = 0$ differ by a factor $\sqrt{2}$. The reason for this is historical and is accounted for in the different estimates of Richardson's constant

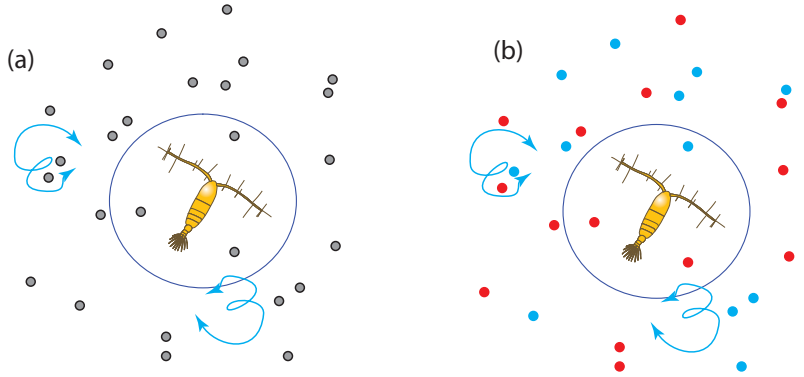


Figure 4.10: The encounter rate of a predator with its prey must be a linear function of the prey concentration. If it is not, then the interrelation of the mechanisms underlying the encounter process would rely on the observer's ability to distinguish between different prey particles. For example, for exactly the same physical situation, an organism's encounter rate with particles of indistinguishable colour would be higher that the sum of its encounter rate with coloured particles.

as

$$w(R) = \alpha_R(\varepsilon R)^{1/3} \tag{4.5.6}$$

The turbulent encounter rate formulation (Eq. 4.5.5) is perhaps less that precise in that it assumes that the all 3 velocities; the swimming of the organisms and its partner, as well a the turbulent velocity, all conform to a Maxwell distribution. While this may be so for the swimming components, it is patently not so for turbulent velocity. However, given the uncertainty in many of the parameters used (α_R for instance, of the precise measure and geometry of R) and the elegance of its form, give Eq. (4.5.5) an certain appeal, and in the absence of specific considerations, this formula gives a reasonably accurate qualitative and quantitative description of turbulent encounter rates.

Turbulent velocity scale: In these expression for turbulent encounter rates (Eq. 4.5.5), I have assumed that the relevant spatial scale for determining the turbulent velocity is the detection distance R, (*i.e.* as specified in Eq. 4.5.6). The derivation above gives the rationale for this. However, for the longest time there was some confusion as to that scale was appropriate. One popular alterative was to use the inter-prey separation distance d; a distance that depends on the concentration of prey C as $d = 0.55C^{-1/3}$ assuming a Poisson prey distribution. However, such a proposition leads to a turbulent encounter rate that is non-linear in prey concentration C and leads to an unsupportable absurdity (Visser and MacKenzie, 1998). Consider a simple example. For a given situation, assuming d is the pertinent length scale (*i.e.* $w = \alpha_R(\varepsilon d)^{1/3}$), the encounter rate is given by $Z = PC^{8/9}$ ($P = \alpha_R\sqrt{2}\varepsilon^{1/3}R^2$ as a convenient constant).

Now consider exactly the same situation, except that now we can distinguish between particles using some biologically and physically nonsignificant parameter, *e.g.* half the particles are coloured blue and the other half coloured red. The total concentration remains the same. The concentration of red particles is $C_r = C/2$, which implies that the predator's encounter rate with red prey items is $Z_r = P(C/2)^{8/9} = 0.54PC^{8/9}$. Likewise, the predator's encounter rate with blue prey items is $Z_b = PC^{8/9}$. The total encounter rate of the predator with

particles of unspecified colour is $Z_r + Z_b = 1.08PC^{8/9}$, which is 8% higher than Z the encounter rate with particles of indeterminate colour. This difference becomes greater the more the prey field can be separated into different species. For instance, 5 different prey items present in equal concentrations would result in an encounter rate 20% higher than if there were only 1 prey item. In other words, the predator's encounter rate depends on the experimenter's ability to distinguish between prey items. Clearly, this is an untenable hypothesis.

Below the Kolmogorov scale: The effect of turbulence on encounter rates is not limited to the inertial sub-range, but penetrates all the way down to molecular scales. This is mediated by turbulent shear, which, in the viscous sub-range is given by

$$\gamma = \left(\frac{\varepsilon}{2\eta}\right)^{1/2} \tag{4.5.7}$$

The flux (*i.e.* encounter rate) to a spherical capture zone is given by

$$Z = \pi R^3 C \left(\frac{\varepsilon}{2\eta}\right)^{1/2} \tag{4.5.8}$$

While this has a somewhat different functional form from that at larger scales ($\varepsilon^{1/2}$ rather than $\varepsilon^{1/3}$, R^3 rather than R^2), encounter rates at scales less than the Kolmogorov scale increase with increasing turbulence.

4.5.1 Turbulence and ingestion: a dome-shaped relationship

Having derived an estimate for turbulent encounter rates (Eq. 4.5.5), there are a few important riders that should be stated. In the first place, encounter rates are not the same as ingestion rate. The general prediction of Eq. 4.5.5 and similar expressions (*e.g.* Rothschild and Osborn, 1988) is that encounter rate is an increasing function of turbulent intensity (or dissipation rate). However, encounter rate is not the same as ingestion rate. When encounter rates are high, ingestion is controlled not by rate of discover, by handling time h. Specifically

$$I = \frac{Z}{1 + hZ} = \frac{C\beta(u, v, R, \varepsilon)}{1 + hC\beta(u, v, R, \varepsilon)} \tag{4.5.9}$$

which is essentially the so-called disk equation (Holling, 1959) leading to a Hollings type II functional response. Increasing turbulence only increases ingestion rate when $hZ < 1$, that is at low prey concentration or low clearance rate or both.

There is a down-side to turbulence: In addition to the saturating effect of handling time, increasing turbulence can have a negative effect on realized ingestion. The detection ability (Saiz and Kiørboe, 1995; Visser, 2001), capture efficiency (MacKenzie and Kiørboe, 2000; Visser and Jonsson, 2000; MacKenzie, 2000) and effectiveness of feeding currents (Yamazaki et al., 1991; Osborn, 1996; Visser and Jonsson, 2000) are all impaired under elevated levels of turbulence.

The dome-shaped response: Thus, in actuality, turbulence can be expected to have a dome-shaped relationship with increasing turbulence – a little is beneficial, but too much is detrimental. This general trend is relatively common in a variety of laboratory studies (Saiz and Kiørboe, 1995; Caparroy et al., 1998; Saiz et al., 2003)).

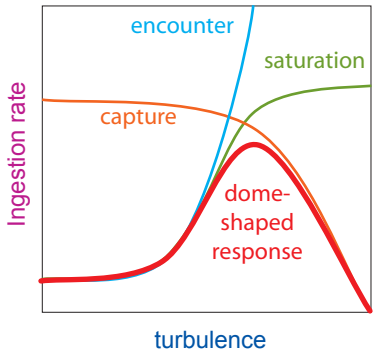


Figure 4.11: In general, while encounter rates increase monotonically with turbulent dissipation rates, ingestion becomes limited first by the saturating effects of handling time (prey simply can't be processed fast enough), and by the negative effects of high turbulence on both detection and capture abilities. The net effect is a dome-shaped relationship.

4.5.2 Vertical positioning and feeding

The dome-shaped response of zooplankton (*e.g.* copepods) feeding is complicated by the vertical structure of turbulence in the water column and the ability of zooplankton species to migrate vertically. Many copepod species exhibit daily vertical migration, avoiding the well-lit surface layers during daylight to reduce predation risk, and returning at dusk to feed (Lampert, 1989). In addition to light and prey abundance, turbulence too varies vertically. Typically, in a stratified continental shelf sea, turbulence levels are high in the surface due to the wind stress, high near the bottom due to tidal and wave induced bottom stresses, and low in pycnocline where vertical density gradients suppress turbulent motion. The possibility therefore exists that vertical migratory behaviour can be used by zooplankton to seek out the depth of turbulence and food concentration that maximizes feeding opportunities and/or minimizes predation risk.

In October 1998, we set out see if these processes indeed operate in nature. The study site was in the northern North Sea where direct turbulence measurements, chlorophyll (*i.e.* an estimate of prey abundance), nutrients, currents as well as detailed measurements of the vertical distribution and recent feeding activity of the copepod community were collected (Visser, 2001).

The recent feeding activity of herbivorous copepods can be deduced from the fluorescence of their gut content relative to the ambient chlorophyll concentration. During the October 1998 cruise, in all 7 groups of herbivorous copepods were in sufficient abundance to be analysed (*Metridia lucens* (females and copepodites), *Centropages typicus* (females and males), *Calanus* spp (females and copepodites), *Pseudocalanus* spp females). All the copepods examined showed a day-night cycle in their feeding activity to a greater or lesser extent, with highest activity at night, and virtually no feeding during the day. Taking this variability into account, the feeding activity can be related to turbulence (Fig. 4.13).

For all the copepod species examined, we found a negative influence of turbulence on the feeding activity; when turbulence was high, the ability of copepods to successfully ingest food was reduced. A similar trend was found (Irigoien et al., 2000) for *C. finmarchicus* in relation to wind speed. This demonstrates at least the downside of the response dome for high levels of turbulence. Turbulent conditions experienced during the October 1998 cruise were at all times high to moderate with dissipation rates seldom falling below 10^{-6} m^2 s^{-3} in the upper 20 m.

With respect to vertical distribution, most species showed little or no diel vertical migra-

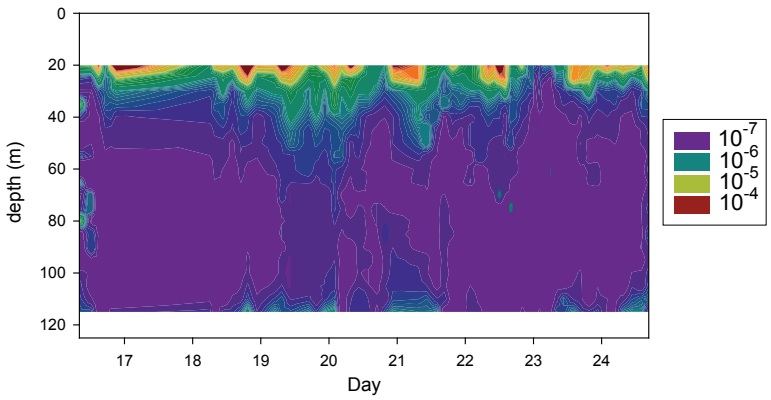


Figure 4.12: The turbulent dissipation rate ($\log_{10}(\varepsilon\ \mathrm{m^2\ s^{-3}})$) as a function of depth during October 1998. Measurements were take from RV Dana at an anchor station in the northern North Sea (1°0′ E, 59°20′ N) in the vicinity of the Fladden Grounds. Data was collected using a Micro Structure-Turbulence (MST) profiler equipped with standard CTD sensors, a microstructure temperature sensor and two microstructure shear sensors (Prandke and Stips 1998). The profiler was used in the free falling mode, sinking at about 0.8 m $\mathrm{s^{-1}}$. The cable was paid out at high speed in order to avoid interfering with the instrument's free fall. Over the duration of the cruise deployments were made every 2-3 h except for a period of ca. 24-h on 17-18 Oct when the sea state was considered too rough for safe deployment.

tion, the exception being *Metridia lucens*. Separating into day and night observations, the relationship of turbulence to the centre of mass of the observed distributions for 7 species separated into female, male and copepodite stages was tested. Results were largely inconclusive except in one case: that of *Oithona similis*. (*O. similis* was by far the most abundant copepod numerically, and so had the best count statistics as well). This species exhibited a significant response to turbulence, avoiding the surface layer at night when turbulence was high. In contrast, during the day, the same species showed no vertical migration in response to turbulence. This difference between day and night indicates that the response is indeed active, and not due to passive mixing. A similar response by *Oithona* off the southern flank of Georges Bank has been reported (Incze et al., 2001). While these observations don't allow us to say precisely how *Oithona* benefits from this behaviour, it may be hypothesized to have to do with how turbulence interferes with *Oithona*'s ability to detect prey.

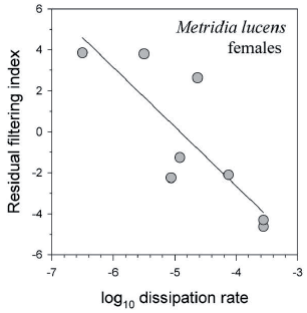


Figure 4.13: Linear regression of the residual filtering index of *Metridia lucens* females, and turbulent dissipation rate in the upper 25 m water column. The residual filtering index is the difference between the measurements, and the best fit diurnal cycle. Only night time observations are included. Turbulent dissipation rate is in $\mathrm{m^2 s^{-3}}$. Fit statistics, $r^2 = 0.69, p < 0.01$ (Visser and Stips, 2002). *Metridia lucens* female exhibit similar behaviour (Visser et al., 2001).

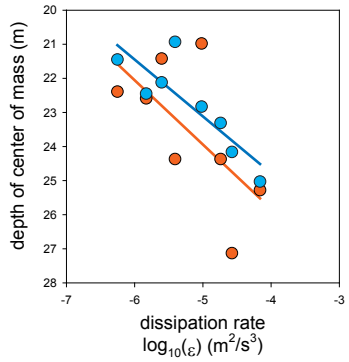


Figure 4.14: Linear regression of the depth of the vertical distribution centre of mass and turbulent dissipation rate for *Oithona similis* females (orange) and copepodites (blue). Only **nighttime** observations are included. Turbulent dissipation rate is averaged over the upper 25 m. Fit statistics, females: $r^2 = 0.40$, $p < 0.05$, copepodites: $r^2 = 0.73$, $p < 0.01$.

4.5.3 Turbulence and feeding currents

Feeding currents are a behavioural adaptation of many zooplankton species (*cf.* Sec. 3.6.4). The interaction of turbulence with a feeding current and its impact of on the encounter rate has been speculated on (Yamazaki et al., 1991; Kiørboe and Saiz, 1995; Saiz and Kiørboe, 1995; Osborn, 1996), but remains largely unknown, although turbulence certainly seems to effect the feeding behaviour of copepods. A relatively simple model to look at the effects on encounter rate is to assume turbulence has a diffusive effect (albeit scale dependent) on an advective feeding current. That is, turbulence has the effect of disrupting the feeding current in a quasi-random way that scales with distance from the current generating mechanism.

The process described in Eq. 4.5.5 (page99) says nothing of the subsequent effect of the encounter on the particle concentration in the vicinity of the encounter sphere. Indeed, it assumes that the concentration remains spatially uniform over time. If we suppose that the outcome of an encounter is to remove a particle, then the encounter rate can be potentially different (Osborn, 1996). In this case, while Eq. 4.5.5 may be adequate to describe the encounter with "new" particles for when the uniform motion of the predator dominates, it can fail when turbulence or random prey motility dominates the relative motion. These latter modes are diffusive in character, so that the particle concentration adjacent to the encounter sphere can be greatly reduced. It is not clear *a priori* what the combined effect of feeding current, spatially varying turbulent diffusion and prey depletion will have on the overall encounter rate.

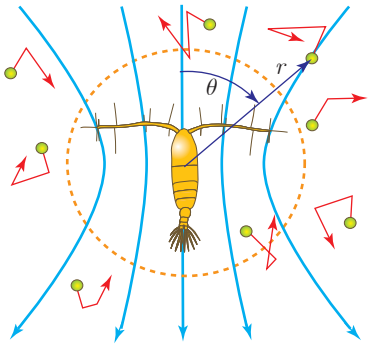


Figure 4.15: Schematic model for the interaction of a copepod feeding current and turbulent diffusion. Spherical polar coordinates (r, θ) are used, and the flow field has rotational symmetry about $\theta = 0$. The copepod has equivalent spherical radius a and its capture zone is the spherical volume radius R.

The effect of turbulence and feeding current on zooplankton feeding are explored using an advection-diffusion model. Specifically, the main features of the model (*cf.* Fig. 4.15) include:

- Spatially dependent turbulent diffusion as expressed in Richardson's law. That is,

$$K_t(r) = \frac{1}{3}\alpha\varepsilon^{1/3}r^{4/3} \approx K_T(r/L)^{4/3} \tag{4.5.10}$$

  where K_T is the turbulent eddy diffusivity.
- Random walk diffusion of prey: Organisms that swim at a constant speed v and randomly change their direction of swimming at an average time interval δ describe a dispersion with characteristic diffusivity

$$K_m = \frac{1}{3}v^2\delta = \frac{1}{3}v\lambda \tag{4.5.11}$$

- Predator feeding current: To first order, a copepod feeding current can be modelled as a stokeslet (Jonsson and Tiselius, 1990), the stream function of which is given by

$$\psi_f(r,\theta) = \frac{3}{4}U_0 a r \sin^2\theta \tag{4.5.12}$$

  where U_0 is the feeding current speed, and a is the equivalent spherical radius of the copepod.

In nondimensionalised form, the advection-diffusion equation governing these processes can be written

$$P\mathbf{u}^* \cdot \nabla^*\phi^* - \left(\kappa + (1-\kappa)r^{*4/3}\right)\nabla^{*2}\phi^* - \frac{4}{3}(1-\kappa)r^{*1/3}\hat{\mathbf{r}} \cdot \nabla^*\phi^* = 0 \tag{4.5.13}$$

where $P = U_0R/(K_T + K_m)$ is the Peclet number[9], $\kappa = K_m/(K_m + K_T)$ is the relative contribution of uniform (Fickian and/or random walk) to total diffusion, $\mathbf{u} = U_0\mathbf{u}^*$ is the flow field relative to the predator, $r = Rr^*$ is the radial distance from the centre of the predator's capture zone, and $\phi = C\phi^*$ is the prey concentration.

The flux entering the capture zone can be calculated as

$$F^* = \int \phi^*\mathbf{u}^* \cdot d\mathbf{r}^* \tag{4.5.14}$$

which gives theoretical results as $P \to 0$ of $F_0^* \to 4\pi/P$ for $\kappa = 1$, and $F_0^* \to 28\pi/3P$ for $\kappa = 0$. On the other hand, as $P \to \infty$, $F_\infty^* \to 3\pi/2$ for any value of κ. Fig. 4.16 shows a series of numerical experiments and converge on the asymptotic values listed above. The transition occurs between $P = 1$ to 10, where the net flux appears to be reasonably well estimated by

$$F^*(\kappa, P) = F_0^*(\kappa, P) + F_\infty^* \tag{4.5.15}$$

[9]Named after the French physicist Jean Claude Eugène Péclet (1793 – 1857). The Peclet number says something about the relative influence of advection to diffusion in determining the local rate of change of some material. It is akin to the Reynolds number; inertia = advection of momentum while viscosity = the diffusion of momentum. It is equivalent to the product of the Reynolds number with the Prandtl number in the case of thermal diffusion, and the product of the Reynolds number with the Schmidt number in the case of mass dispersion. The ratio of the Schmidt number to the Prandtl number is the Lewis number. A final dimensionless number of some relevance is the Sherwood number; the ratio of advective to diffusive *flux*, and while in general it is a function of the Peclet number, it can be quite different particularly in the vicinity of solid boundaries and absorbing/exuding bodies, *i.e.* sources and sinks of momentum and diffusive material.

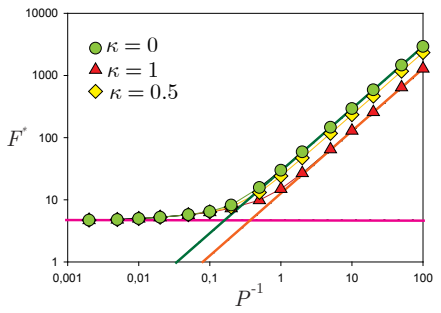


Figure 4.16: The non-dimensionalised flux of prey particles into the copepods capture zone as a function of inverse Peclet number, P^{-1}.

In a practical sense, Fig. 4.16 shows that the total flux of prey into the capture zone of a feeding current generating predator in the presence of turbulence, is given by the sum of the fluxes from the component processes. That is, for a pure feeding current of strength U, the flux is

$$Z_{fc} = \frac{3}{2}\pi C R^2 U, \tag{4.5.16}$$

while for a passive predator feeding on a prey describing a random walk, the flux is

$$Z_{rw} = \frac{4}{3}\pi C R v \lambda. \tag{4.5.17}$$

The contribution from turbulence for a passive prey and predator is

$$Z_{turb} = \frac{28}{9}\pi C (\alpha\varepsilon)^{1/3} R^{7/3}. \tag{4.5.18}$$

The total flux for any combination of these is well approximated by

$$Z_{tot} = Z_{fc} + Z_{rw} + Z_{turb}, \tag{4.5.19}$$

justifying the idea of the addition of encounter kernels (Kiørboe and Saiz, 1995). Furthermore, the simulations demonstrate that the combined effect is only significant for Peclet numbers P between 1 and 10. For $P > 10$, advection dominates and $Z_{tot} \approx Z_{fc}$, while for $P < 1$, diffusion dominates and $Z_{tot} \approx Z_{rw} + Z_{turb}$. The flux (*i.e.* encounter rate) given by these expressions have the same scalar dependence as those given for each process in Rothschild and Osborn's theory, Eq. 4.5.5 (page99), although they differ in some details. This difference arises essential because Rothschild and Osborn's theory is formulated for an invariant prey concentration (*i.e.* a ballistic rather than a diffusive scenario) whereas we consider the steady state situation when the prey concentration has adjusted to the depletion rate imposed by feeding. The time scale for this adjustment is of the order

$$T \approx \frac{R^2}{K_m + K_T}, \tag{4.5.20}$$

e.g. 1 to 30 seconds for a typical copepod in a shelf sea. The actual situation found in nature would lie between these two extremes, and would depend to the rate at which the predator jumped to a "new" prey field.

4.5.4 Turbulence and signal detection

The question now arises, how does turbulence effect detection by hydromechanical signals. Lets consider a simple example. A copepod with inherent sensitivity s_0, can detect a swimming ciliate at a distance $R_0 = a(U/s_0)^{1/2}$ (Svensen and Kiørboe, 2000). The effect of turbulence is to introduce velocity fluctuations between the body of the copepod, and the setae positioned along its antennae. Given the length of each antenna to be b, the magnitude of the turbulent velocity fluctuations at the antenna tips is $w = \alpha(\varepsilon b)^{1/3}$ from Eq. 4.5.6 (page 100). For detection, one can assume that the ratio of the signal (from a swimming ciliate) to noise (turbulence) must exceed some level κ. With this constraint, the detection distance as a function of turbulent dissipation rate to follow $R(\varepsilon) \propto \varepsilon^{1/6}$ can be predicted; a prediction reasonably well supported by observations (Saiz and Kiørboe, 1995; Visser, 2001).

While hydromechanical signals provide a feasible means by which zooplankton can gain information on their neighbours, these signals must be "overheard" above the background turbulent "noise". The functional form of the detection distance means that the detection of different modes of motility (sinking, swimming, feeding current) is susceptible to turbulent interference to differing degrees.

4.5.5 Turbulence, behaviour and prey selection

In the planktonic arms race, the ability of prey to detect and escape from an attacking predator (Caparroy et al., 2000) and the ability of a predator to approach a prey without alerting it (Viitasalo et al., 1998) appear to be important factors in the dynamics and structure of planktonic communities (Greene, 1983). Some copepods change their mode of feeding from suspension feeding to ambush-feeding depending on the type of prey available (Saiz and Kiørboe, 1995). For example, *Acartia tonsa* switches between suspension feeding on diatoms, and ambush feeding on ciliates. This appears to be due to the ability of ciliates to detect and escape from copepod feeding currents balanced against the fact that swimming ciliates have a higher encounter probability and are hydromechanically more conspicuous than immotile diatoms. The effect of turbulence on ingestion is negative for both feeding current and ambush feeding although its effect on ambush feeding is more acute.

A risk - benefit balance appears to be in effect in many copepod and nauplii escape reactions. Copepods, in detecting an approaching predator or feeding current, can perform spectacular escape jumps (Fields and Yen, 1996a). All things being equal, the distance at which they can detect a predator increases with their size (Kiørboe and Visser, 1999). While the escape jump can remove the prey from proximate danger, it also has an element of risk in that the escape jump makes the prey hydrodynamically more conspicuous to a rheotactic predator (Landry, 1980; Yen and Strickler, 1996; Tiselius et al., 1997)as well as visually more conspicuous to a visual predator like a fish larvae (Viitasalo et al., 1998). Modelling the detection – escape ability of different copepod life stages has been successfully used to describe observed size selectivity of different predators (Caparroy et al., 2000). Although never properly explored, impaired detection by turbulent interference will tend to expose smaller prey to greater risk. The question of the risk-benefit-cost tradeoff will be given more detailed examination in an upcoming section (Sec. 7.3.3).

4.6 Sinking through turbulence

One aspect of particle-flow interactions that may seem counter intuitive, is that particles sinking in turbulence actually increase their average sinking speed as compared to that in

still water (Fig. 4.17).

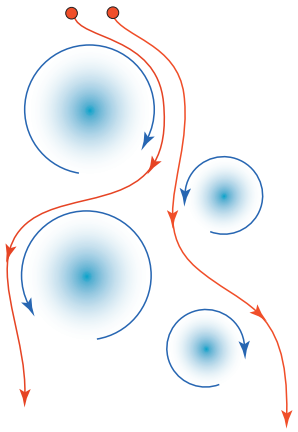


Figure 4.17: Turbulence increases sinking speed. This may appear puzzling as turbulence (mixing and stirring) keep particles (sediment and plankton in the ocean, tea-leaves in a more a more day-to-day setting) in suspension longer. A negatively buoyant particle sinking through a field of eddies, will preferentially be swept into the down-flowing side of eddies. If we look at the possible paths a particle takes while sinking through a field of eddies, it becomes clear that particles will experience more down-flows than up-flows.

We can examine this more closely by integrating the simplified particle equation (Eq. 3.2.39) for a negatively buoyant particle $\alpha > 1$ in an idealized, horizontally inclined vortex. Results for such a model are presented in Fig. 4.18. In general most particles sink faster than would be expected in the absence of the vortex. However, some of the particles become trapped in what might be called the "tumble-drier" effect; they tumble in orbits being alternately swept upward by the eddy and sinking downward, and eventually spiralling in to a point. In nature however, vortices are ephemeral, so the trapping effect would have little impact in the long term.

One final, curious effect for sinking particles is their reduced dispersion. That is, a cluster of particles sinking sinking at the same rate through a turbulent water column will have a lower rate of dispersion (*i.e.* their rate of increase of variance will be less) than a similar cluster of neutrally buoyant particles.

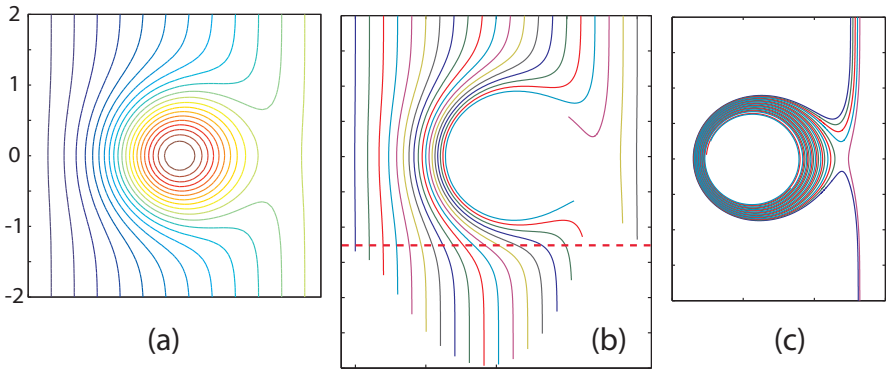


Figure 4.18: Sinking of negatively buoyant particles through an eddy. (a) stream function of the eddy, (b) trajectories of particles, (c) detailed trajectory of particles that become trapped. The horizontal dashed red line in (b) indicated the depth to which particles would sink in the absence of the eddy.

4.7 Chemical trails in turbulence

Although seemingly uniform, the aquatic environment in which plankton live is filled with ephemeral chemical and physical cues that have scales of meters and smaller. Chemical plumes trail behind falling particles and swimming animals; small velocity perturbations follow swimming or feeding copepods. The ability of an organism to survive depends on its use of these signals to find food and mates, and to avoid predation; our understanding of the planktonic ecosystem depends on our ability to understand and describe the heterogeneity of this world.

Given the potential of chemicals as a source of environmentally relevant information, it is not surprising that nearly all zooplankton have developed chemo-sensory abilities. Bacteria climb chemical gradients to find choice resources, male copepods trail pheromone trails to find females, chemo-sensory organs near mouth parts test whether algal cells should be ingested or discarded, parasites chemically track their hosts, and organisms show altered behaviour when exposed to exudates from their predators.

In the oceans, the physics relevant to particulate organic matter includes the control of settling and associated water movement, and the subsequent diffusive spread of released molecules. Organic matter released by falling particles can provide localized food resources for microbes in the surrounding water column and provide chemical cues for animals and microbes to locate the particles. Copepods, for instance, respond to dissolved organic material in the water and can follow a chemical trail to a sinking aggregate or to other animals (Poulet and Ouellet, 1982; Kiørboe, 2001). The detection of a specific particle depends on the rate at which the particle releases material, its sinking rate, the chemical nature and the molecular diffusivity of the chemical substrate, the threshold concentration detectable by an organism, and the geometric patterns these threshold concentrations trace out in 3D space for variable conditions.

Plumes offer significantly different resources for osmotrophs and for hunters. For bacteria or phytoplankton located in the enriched environment behind a leaky detrital particle or zooplankter, the plume volume defines the favourable region. Used together with information on particle or animal distributions, it can be used to describe the heterogeneity of the chemical environment (Jackson and Kiørboe, 2003). To understand zooplankton search behaviour, plume length and cross-sectional area relate to the rate of finding particulate targets (Jackson and Kiørboe, 2003). The detectable plume behind a large falling particle in *still water* could be as long as 1 m (Kiørboe and Thygesen, 2001), but what is the effect of turbulence in lengthening or in disrupting the plume?

4.7.1 Modelling a chemical trail in turbulence

Describing the characteristics of a chemical plume trailing behind a sinking detrital particle is a relatively generic problem in the marine pelagic environment. Such detrital particles represent a rich resource for those animals that can find them, both in the organic carbon the particles themselves represent, but also the community of fellow-travellers – bacteria, protists and zooplankton – these particles attract. Furthermore, detrital particles are considered a major vehicle for the vertical transport of organic carbon from the euphotic zone to the deep ocean. How, where and to what degree these particles are colonized and remineralized on their downward odyssey has direct implications for the efficiency of the biological pump, and the sequestering of organic carbon in the deep ocean. Consider for instance, a horizontally swimming chemo-sensitive copepod (such as the harpacticoid *Microsetella*), searching for a

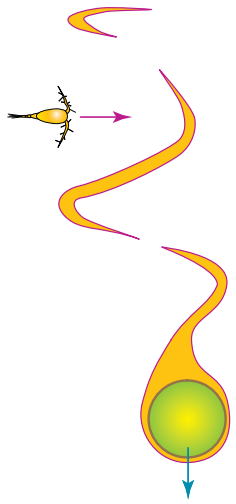


Figure 4.19: Sinking detrital material, such as faecal pellets, marine snow and discarded appendicularian houses, leave a trail of organic solutes that chemosensory copepods can track to their source Kiørboe (2001). Swimming horizontally a searching organism will encounter the trail at a rate proportional to its swimming speed, and the cross sectional area of the trail – defined by the organism's detection threshold – projected on the vertical plane. The effect of turbulence is to elongate the plume and possibly break it into segments.

vertically sinking detrital particle exuding some dissolved organic carbon compound (Fig. 4.19). The basic physical properties that prescribe the problem depend on the properties of the particle those of the searcher, as well as those of the environment:

- Particle: 2 basic properties are set particle size, r, sinking speed w and the rate at which they exude chemicals Q. For instance, the settling rates of marine snow and fecal pellets have been described using a power law dependence on particle size.[10] Similar power law relationships exist for the organic content of particulate matter, and for the rate at which amino acids and other dissolved organic compounds are released (Kiørboe and Thygesen, 2001).
- Searcher: The concentration at which an organism can detect the plume, C^*, determines the effective extent of a plume. Its value is clearly specific to organism and compound. Estimates have been made using observed background concentrations, and observed responses by organisms (Kiørboe and Thygesen, 2001; Jackson and Kiørboe, 2003) giving estimated values of 3 to 4×10^{-11} mol cm^{-3}. The other searcher property is its swimming speed and behaviour.
- Environment: The diffusivity D of the chemical compound has immediate implications for the plume characteristics. In addition, if the fluid motion, not just turbulence, but also laminar shear, will effect the plume.

In still water, the length of the plume can be calculated as

$$Z_0 = \frac{Q}{4\pi D C^*} \tag{4.7.1}$$

Somewhat surprisingly, the length of the plume does not depend on the speed at which the particle sinks; sinking faster reduces the initial concentration of released material in such a way that the two effects compensate precisely. Other relevant measures of the plume are its volume V_0 and cross-sectional area σ_0; the latter particularly so for horizontally swimming copepod searchers (*cf.* Fig. 4.19).

[10] Theoretically, small particles should follow Stokes' law with $w \propto a^2$. An observed power law with exponent other than 2 does not invalidate Stokes' law, but reflects an imprecision in the concept of particle size and density. Marine snow particles are aggregates with variable fractions of "empty space" changing the effective density and/or linear dimension.

In turbulent water, the plume will be stretched, increasing gradients of the chemical concentration and thus the flux out of the plume. Furthermore, this rate of stretching will be uneven along the length of the plume, breaking it into sections. In order to simulate these effects we use a kinematic simulation (*cf.* Sec. 4.3.1) through which a point source, advected by the flow while sinking, lays out a chemical trail. To make the problem tractable, the plume is considered a sequence of small cylindrically symmetric Gaussian distributions of chemical tracer

$$\phi_{i,j}(\varrho) = C_{i,j} \exp(-\varrho^2/\lambda_{i,j}^2) \tag{4.7.2}$$

that are stretched by turbulence according to the kinematic simulation, and diffuse by molecular diffusion. Here, $C_{i,j}$ is the chemical concentration along the axis of segment j of the plume at time t_i. Likewise, $\lambda_{i,j}$ is the e-folding scale of the Gaussian distribution in this section at the same time. This approach allows a fairly full and parsimonious description of the evolving characteristics of the plume (its length, cross-sectional area and volume) since both stretching (or squashing) and diffusion transform one Gaussian distribution into another. That is, after a small time step δ, the distribution in Eq. (4.7.2) becomes

$$\phi_{i+1,j}(\varrho) = C_{i,j} \frac{\lambda_{i,j}^2 s_{i,j}}{4D\delta + \lambda_{i,j}^2 s_{i,j}} \exp\left(\frac{-\varrho^2}{4D\delta + \lambda_{i,j}^2 s_{i,j}}\right) \tag{4.7.3}$$

where $\ell_{i,j}$ is the length of section j at time $t = i\delta$, and $s_{i,j} = \ell_{i,j}/\ell_{i+1,j}$ gives a measure of the stretching of this section over consecutive time steps. Thus the plume characteristics can be derived from the updating process:

$$\lambda_{i+1,j}^2 = 4D\delta + \lambda_{i,j}^2 \frac{\ell_{i,j}}{\ell_{i+1,j}} \tag{4.7.4}$$

$$C_{i+1,j} = C_{i,j} \frac{\lambda_{i,j}^2}{\lambda_{i+1,j}^2} \frac{\ell_{i,j}}{\ell_{i+1,j}} \tag{4.7.5}$$

which are closed by the initial conditions

$$\lambda_{1,j}^2 = 4D\delta \tag{4.7.6}$$

$$C_{1,j} = \frac{Q}{4\pi D w \delta} \tag{4.7.7}$$

4.7.2 Trail characteristics and metrics

Armed with these simulation tools, it is a fairly straightforward to estimate the plume characteristics for different particle sizes, leaking rates, detection levels and turbulent conditions. Animations give striking illustrations of the stretching, dissipation and pinching off of plumes. Yet despite the complexity of the governing processes as well as the large parameter space involved, the results collapse into a few seemingly universal curves.

Total plume cross-sectional area σ^* and volume V^* for instance, when normalized against the same metric in still water, show a remarkably consistent pattern with respect to normalized turbulence. Specifically, these appear to follow functions of the form

$$\frac{\sigma^*}{\sigma_0^*} = \frac{1}{1 + a\gamma T_0^*} \tag{4.7.8}$$

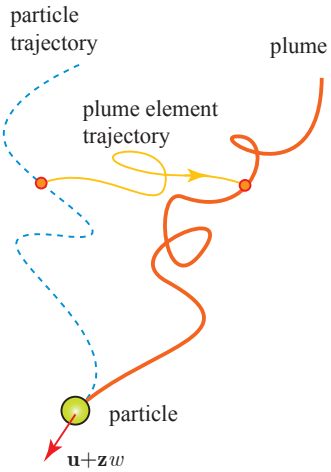


Figure 4.20: The plume construct. A particle sinks with settling speed w while being advected by a turbulent flow field $\mathbf{u}$. Neutrally buoyant tracers are released from the particle at discrete time intervals and subsequently tracked through the turbulent flow field. A snap-shot of these tracers at a given time traces out the centre line of the plume. The concentration within the plume can be estimated by calculating the combined effects of turbulent stretching and diffusion over time.

(*cf.* Fig. 4.21) where a is a fit parameter, $\gamma = \sqrt{\varepsilon/6\eta}$ is the turbulent shear, and T_0^* is the inherent plume time scale in still water. That is

$$T_0^* = \frac{Z_0^*}{w} = \frac{Q}{4\pi D w C^*} \tag{4.7.9}$$

which is the time after release it takes for the concentration in a small section of the plume to fall below the detection threshold. While both volume and cross-section area have this functional form, the normalized plume length Z^*/Z_0^* appears to remain invariant with respect to normalized turbulent shear.

A simple theoretical argument explains these results. In 3D isotropic turbulence, the average

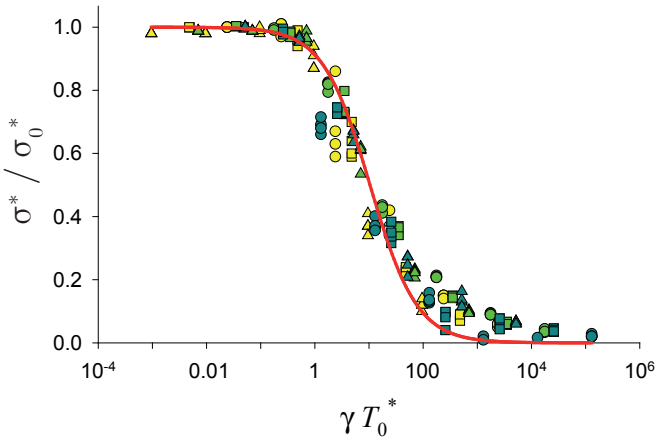


Figure 4.21: The normalized cross-sectional area on a vertical plane of a plume as a function of normalized turbulent shear. Different symbols and different colours represent a 3 different sinking speeds, and 3 different detection thresholds. 5 turbulent intensities were run on each combination in 4 replicates giving a total of 180 model runs. A parametric fit of the form Eq. 4.7.8 corresponds to a fit parameter $a = 0.1 \pm 0.01$ for marine snow and fecal pellets (Visser and Jackson, 2004).

distance between 2 parcels in the fluid changes at a rate that is scale-dependent. In the inertial sub-range (scales larger than the Kolmogorov length scale ζ and smaller than the integral length scale L), this is given by Richardsons law:

$$\frac{\partial \ell}{\partial t} = \alpha(\varepsilon \ell)^{1/3} \tag{4.7.10}$$

where ℓ describes the mean separation between any 2 points in the fluid. However, the distance between these 2 points following a *material line* could be considerably longer and is governed by the micro-structure rate of strain $\gamma_T = [\varepsilon/(2\eta)]^{1/2}$ associated with the turbulent flow (*e.g.* Batchelor, 1952; Cocke, 1969; Monin and Yaglom, 1975; Tsinober, 2001). A material line is a mathematical construct and is the line traced out by infinitesimally small fluid elements at some moment in time. As the fluid flows, the fluid elements move and the material line stretches, folds and deforms.

At sub-Kolmogorov length scales, the relative distance between 2 fluid elements is given by:

$$\frac{\partial d\xi}{\partial t} = \gamma d\xi \tag{4.7.11}$$

where $d\xi$ is the length of an infinitesimally small material line segment and γ is the turbulent rate of strain associated with material line deformation (Monin and Yaglom, 1975). Assuming that γ is isotropic, the length of a material line segment is thus given by: $d\xi(t) = d\xi_0 e^{\gamma t}$. For a material line being traced out by a point source sinking at a speed w, the initial segment length is given by $d\xi_0 = wdt$, where dt is a small time step. The total length of the material line traced out by the falling particle is given by integrating over the path through time:

$$\xi(t) = (e^{\gamma t} - 1)w/\gamma \tag{4.7.12}$$

This gives the length of a material line traced out by a point moving in a uniform direction at speed w under isotropic turbulent straining. t here can be interpreted as the "age" of a fluid element since it left the source.

By induction, the concentration along the centre line of the plume can be written as

$$C(t) = \frac{Q\gamma}{4\pi D w}\frac{1}{e^{\gamma t} - 1} = C^* T_0^* \frac{\gamma}{e^{\gamma t} - 1} \tag{4.7.13}$$

from which it can be readily deduced that the time after which a plume segment becomes undetectable is given by

$$T^* = \frac{1}{\gamma}\ln(\gamma T_0^* + 1) \tag{4.7.14}$$

The length of the plume at this time is simply

$$Z^* = \xi(T^*) = \frac{Q}{4\pi D C^*} = Z_0^* \tag{4.7.15}$$

that is, equivalent to the plume length in still water; a prediction borne out in the numerical experiments.

The radial distance at which the plume can be detected is given by

$$(\varrho^*)^2 = \frac{4\pi\xi}{\xi\gamma + w}\ln\left(\frac{Z_0^*}{\xi}\right) \tag{4.7.16}$$

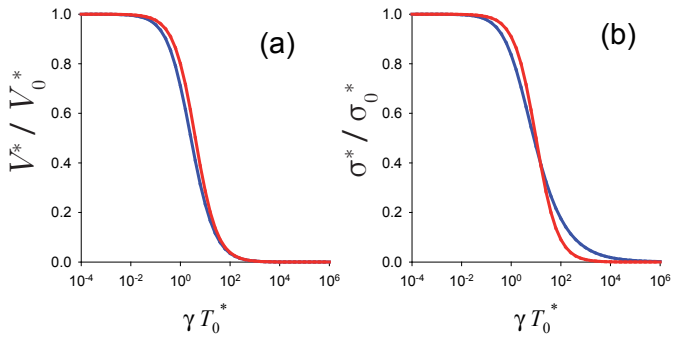


Figure 4.22: The normalized volume (a) and cross-sectional area (b) as a function of normalized turbulent shear. Comparison of the theoretical curves, and the empirical fit to numerical experiments

The volume of the plume is thus given by:

$$V^* = \pi \int_0^{Z^*} (\varrho^*(\xi))^2 \, d\xi = 4V_0^* \int_0^1 \frac{\xi'}{\gamma T_0^* + \xi'} \ln(\xi') d\xi' \tag{4.7.17}$$

where $\xi' = \xi / Z_0^*$. The integral does not yield an analytic solution, although it can be numerically solved. In a similar manner, the cross sectional area can be computed as:

$$\sigma^* = 2 \int_0^{Z^*} \varrho^*(\xi) d\xi = \sigma_0^* \int_0^1 \left(\frac{-\xi' \ln(\xi')}{\gamma T_0^* + \xi'} \right)^{1/2} d\xi' \tag{4.7.18}$$

Numerical evaluations of the theoretical predictions from Eq. (4.7.17) and Eq. (4.7.18)) show that they are essentially the same as those calculated by fitting Eq. (4.7.8) to the result of the simulations for values of a representative of marine snow and fecal pellets (Visser and Jackson, 2004). This is a rather remarkable result – that the complex stretching and diffusion of material exuded from a travelling sources should reduce to a near universal phenomenology, functionally dependent on a handful of parameters.

4.7.3 Appendicularians and Microsetella

Following form these results, we thus have a relatively robust estimate of encounter relevant measures of the chemical trails under turbulent conditions. In particular, since turbulence is a function of depth, we have a means of estimating how fast, where and to what degree sinking detrital material is remineralized in the water column. Such questions have an important role in our understanding of the biological pump, the rate at which organic carbon transported to the deep ocean. As an illustrative example, I present the case of discarded appendicularian houses and the harpacticoid copepod *Microsetella norwegica*.

Appendicularians, also known as larvaceans, are common members of many marine pelagic ecosystems and can reach abundances of 300 m^{-3} in the upper reaches of the water column (Vargas et al., 2002). They construct a mucus housing through which they filter seawater. The filters become clogged with time, forcing animals to abandon the old and build new houses, typically at a prodigious rate, 3 to 4 times a day (Hansen et al., 1996). The abandoned houses can form a significant fraction of the particulate organic matter settling out of the surface layer, not only in themselves and the organisms and detritus they contain, but also the material and organisms that stick to them while sinking. Discarded, sinking appendicularian houses and the aggregates they form are mini-ecosystems in their own right, exchanging organisms, particulate and dissolved material with their ever-changing fluid environment. The harpacticoid copepod *Microsetella norwegica* is apparently an important colonizer and consumer of detrital aggregates in temperate shelf seas. Further, it appears that *Microsetella norwegica* uses remote chemical detection since theoretical computations suggest that this encounter mechanism is the only means by which sufficient encounters can be achieved to account for the reported high clearance of aggregates (Kiørboe, 2000; Kiørboe and Thygesen, 2001; Koski et al., 2005).

The cross sectional area σ of a chemical plume behind a sinking appendicularian house decreases with increasing turbulence (Fig. 4.23). For an *Oikopleura dioica* house sinking at around 10 m d^{-1} (Maar et al., 2004), the cross section of the plume is 10 times greater at relatively low turbulence ($\varepsilon = 10^{-8}$ m^2 s^{-3}) than it is at higher turbulence levels (10^{-6} m^2 s^{-3}). This is consistent with Maar et al. (2006), *i.e.* the dissipation range at which *Microsetella norwegica* seem to avoid turbulence at levels exceeding 10^{-7} m^2 s^{-3}.

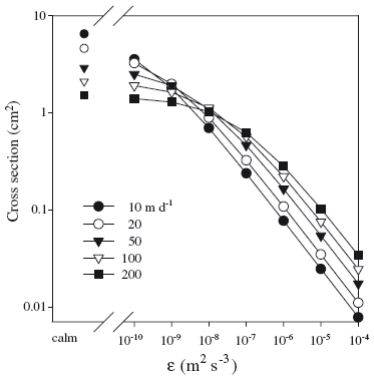


Figure 4.23: Cross sectional area of the chemical plume versus turbulent dissipation rate ε for marine snow aggregates with a fixed particle radius of 0.2 cm and different sinking velocities (10, 20, 50, 100 and 200 m d^{-1}). The cross sections were calculated according to Eq. 4.7.8 with $a = 0.1$ (Visser and Jackson, 2004) and assuming a leakage rate of 9×10^{-14} mol s^{-1} and a threshold detection concentration of 3×10^{-8} μM (Jackson and Kiørboe, 2003).

A general trend that can be deduced from Fig. 4.23 is that for high turbulence levels $\sigma \propto \varepsilon^{-1/2}$. That is, for ever 2 decade of turbulence increase, the cross sectional area of the plume decreases by a decade.

The rate at which copepods encounter discarded houses constrains the rate at which the animals can consume the houses. A copepod swimming horizontally at a speed v has an encounter kernel β, *i.e.* maximum clearance rate, given by: $\beta = v(\sigma + 2b\ell)$ where σ and ℓ are the plume cross-section and length respectively, and b is the length of the copepod's antenna (along which its chemosensory setae are arrayed). For copepods associated with marine snow aggregates (*e.g. Microsetella* spp., *Oncaea* spp.), a representative cruising speed is $v = 1$ mm s^{-1} and a representative antenna length is $b = 100$ μm.

Using these parameters, the clearance rate β varies from 25 l d^1 in perfectly still water, to

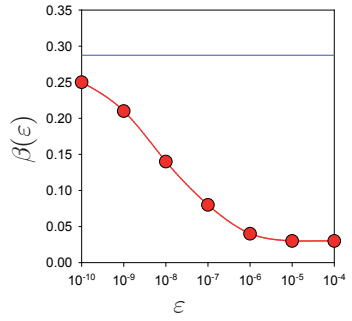


Figure 4.24: Clearance of a chemosensory, horizontally cruising copepod (antenna length 100 μm, swimming speed 0.1 cm s.1, detection threshold 3×10^{-8} μM) on sinking appendicularian houses (size 2 mm, leaking rate 9×10^{-14} mol s^{-1}) as a function of turbulent dissipation rate. The horizontal dotted line indicates clearance rate in perfectly still water, while the vertical dashed line indicates the suggested turbulence threshold for triggering an avoidance response

about 3 l d^1 at turbulence levels ($\varepsilon > 10^{-6}$ m^3 s^2), typical for instance within the surface mixed layer during strong wind events (*cf.* Fig. 4.24). This would suggest a potential trade-off for chemo-sensitive, trail seeking copepods; they should choose a depth where turbulence is sufficiently low to allow efficient plume detection, but not too deep so as to be out-competed by competitors higher in the water column. This also has implications for the depth at which houses are remineralized, and the ensuing flux of carbon to deep waters. Taking $\varepsilon = 10^{-7}$ m^2 s^{-3} as a preferred turbulence level would suggest that appendicularian houses are fed on by copepods at depths typically below the pycnocline. While some of the material may be brought back to the surface to be respired and excreted by these copepods as they migrate vertically, considerable amounts will likely remain to be remineralized at depth.

4.8 Biomixing of the oceans

Biomixing, the action of swimming organisms in mixing the worlds oceans, is a concept that captured attention recently. If correct, biomixing has far-reaching consequences for our understanding of the oceans. For instance, at the global scale, it suggests that billions of small organisms paddling away in the deep oceans stir cold deep water upward (Dewar et al., 2006), thus contributing to global circulation and ultimately impacting global climate. At more local scales, it also suggests that schools of krill and other marine animals "plough" the thermocline (Huntley and Zhou, 2004), mixing nutrient-rich water upward and thereby fertilizing their own feeding grounds.

Swimming organisms do seem to dissipate substantial amounts of mechanical energy. There are even observations showing considerably elevated dissipation rates in the wake of a migrating school of krill (Kunze et al., 2006). The case for biomixing thus seems to be compelling.

But can swimming organisms actually achieve significant mixing? Central to this question is the mixing efficiency involved. In calculating the impact of biomixing, the dissipation of mechanical energy has been *equated* with mixing. Yet, most of the biomixing is purportedly achieved by small but numerous zooplankton of 1 cm size or less. Can mechanical energy at these small scales achieve any substantial mixing (that is, increase the potential energy of the water column, *cf.* Fig. 4.25) before it is dissipated as heat?

Turbulence in the oceans is generated by a variety of mechanisms, including tides, winds, as well as swimming animals. It cascades energy from large scales to ever smaller scales, where it is eventually dissipated. Turbulence is effective in mixing because it is active over a range of scales; stretching and folding of the fluid at large scales facilitates mixing by molecular diffusion at smaller scales. That is, dissipation rate is not the only aspect of turbulence involved in mixing (*cf.* Fig. 4.26).

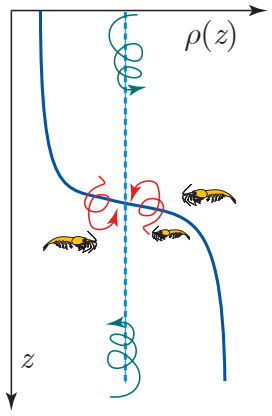


Figure 4.25: Typically, the world's oceans are stratified. In the deep ocean basins, cold salty waters produced in polar seas flood the deep, to be replaced by surface waters warmed in equatorial and tropical regions flowing poleward. In temperate regions, this is augmented by the seasonal heat and cooling of surface waters. The potential energy of a well-mixed water column (density profile given by the dashed line) is greater than that of a stratified water column (solid line). That is, work has to be expended for mixing to take place. Winds, tides, internal waves as well as swimming organisms are potential sources of this energy.

The efficiency of turbulence in mixing a stratified water column is expressed by Γ, the ratio of the change in potential energy to the work done (Fig. 4.27). Mixing efficiency is controlled by three parameters: the integral length scale L (the scale at which turbulent kinetic energy is imparted to the flow, and therefore the scale at which most of the turbulent kinetic energy is found), the rate of turbulent energy dissipation ε (equivalent to the rate of work done), and the Brunt-Väisälä (or buoyancy) frequency N given by

$$N^2 = -\frac{g}{\rho_0}\frac{\partial \rho}{\partial z} \tag{4.8.1}$$

as a measure of the stratification of the water column.

The rate at which potential energy increases due to turbulent mixing is KN^2, where K is the

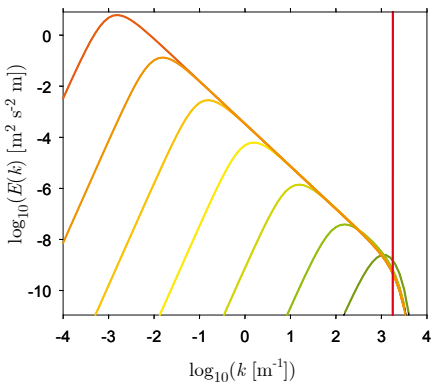


Figure 4.26: While the turbulent energy spectrum has a universal shape considered to be dependent only on 2 parameters: ε energy input rate (equal to turbulent dissipation rate) and μ the fluid viscosity, actual realizations also depend on L the integral length scale, the scale at which kinetic energy is imparted to the fluid. Sketched here are the turbulent energy spectra for a series os realizations all where ε and μ are equivalent but where L varies over several orders of magnitude; from basin scale to the size of an individual plankter. It would be naïve to suppose that all of these turbulent energy spectra would have the same mixing characteristics.

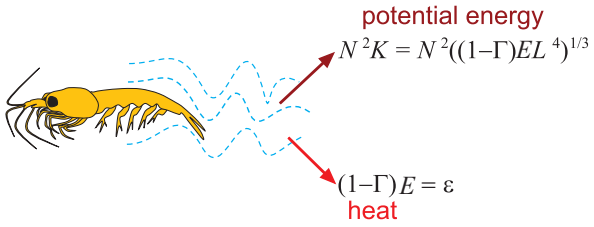


Figure 4.27: The partition of mechanical energy imparted by a swimming organism to a fluid: a portion is lost a heat through viscosity, and the remaining portion can go into mixing, *i.e.* increasing the potential energy of the water column.

effective diapycnal turbulent diffusivity. If E is the total rate at which energy is imparted to the fluid (and thus also the total rate at which it is dissipated), then the mixing efficiency is given by $\Gamma = KN^2/E$. This is the fraction of the total rate energy of energy loss going into changing potential energy. The remaining fraction dissipates as heat through viscosity acting on microstructure shear. That is $\varepsilon = (1-\Gamma)E$. The total energy balance thus requires that

$$E = N^2K + (1-\Gamma)E \tag{4.8.2}$$

The point to note is that turbulent diffusion K is in itself also a sink of energy. That is, in order to diffuse, fluid elements (and thus any tracers they contain) have to move relative to each other, and thus work has to be done against viscosity.

A phenomenological relationship between dispersion and the rate of work done comes from Richardson's law (*cf.* Sec. 4.1.2), giving

$$K_L \approx \alpha(\varepsilon L^4)^{1/3}$$

where L is the integral length scale - the scale of the energy containing turbulent eddies. The rate of dispersion by turbulent motion and turbulent diffusivity are not identical. However, they are not entirely unrelated either. The rate of dispersion is equivalent to the turbulent diffusivity for material with vanishingly small molecular diffusivity. $K_L \approx K_H + D$ where K_H is the turbulent diffusivity in an unstratified fluid, and D is the molecular diffusivity. While there is no direct phenomenological relationship between dispersion rate and diapycnal diffusivity K, we can state with some certainty that

$$K_L \approx K_H > K$$

This is because in a stratified fluid, some of the energy that would have gone into dispersing fluid elements in the unstratified case, may be radiated away through internal waves. Taking $K = K_L$ as an upper limit of diapycnal diffusivity, and thus corresponding to an upper limit to mixing efficiency, Eq. 4.8.2 becomes

$$E = \alpha((1-\Gamma)EL^4)^{1/3}N^2 + (1-\Gamma)E$$

which can be solved for Γ to give

$$\frac{\Gamma}{1-\Gamma} = \alpha\left(\frac{L}{B}\right)^{4/3} \tag{4.8.3}$$

(*cf.* Fig. 4.28) where, B is the buoyancy length scale given by

$$B = \left(\frac{(1-\Gamma)E}{N^3}\right)^{1/2} = \left(\frac{\varepsilon}{N^3}\right)^{1/2} \tag{4.8.4}$$

The asymptotic limits of Γ are

$$\Gamma \approx \Gamma_0 \quad \text{for} \quad L \geq B \tag{4.8.5}$$

$$\Gamma \approx (L/B)^{4/3} \quad \text{for} \quad L < B \tag{4.8.6}$$

Theoretically, Γ_0 can approach 1, although observations indicate that it is close to 0.2.

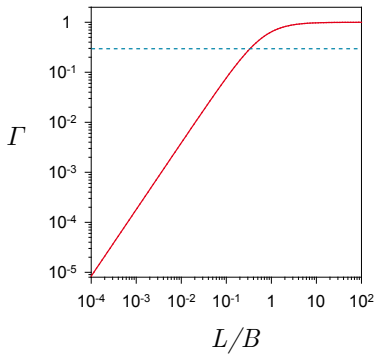


Figure 4.28: The turbulent mixing efficiency Γ and a function of the ratio of the integral length scale L and the buoyancy length scale B.

The net dissipation rate due to an assemblage of swimming organisms depends on the power expended per individual and the number of individuals per unit volume (Huntley and Zhou, 2004). Thus, the dissipation rate ε of a school of krill (assuming a body length of 1 to 1.5 cm, a swimming speed of 5 to 10 cm/s, and a number density of 5000 individuals m^{-3}) is equal to 10^{-5} to 10^{-4} W/kg, consistent with observations (Kunze et al., 2006).

How much mixing does this represent? An organism of a given body size ℓ cannot inject energy into a flow at length scales larger than itself. Thus $L \approx \ell$, consistent with observations for grid-generated turbulence. The buoyancy frequency for the surface ocean is typically 10^{-2} s^{-1} or less, so that the buoyancy length scale associated with the above measurements is 3 to 10 m, and the corresponding mixing efficiency $\Gamma = 10^{-4}$ to 10^{-2}. Hence, only 1% at most of the mechanical energy dissipated by the swimming school of krill and other marine animals actually goes into mixing. The dissipation rate measured in the wake of a dense assemblage of swimming organisms may indeed be considerably higher than that associated with oceanic turbulence, but it does not necessarily follow that the corresponding mixing is also proportionally higher.

The case for biomixing as an important component of the meridional overturning circulation is fraught with the same problem. Considering tides and winds alone, there is an apparent shortfall of ~ 1 TW (*i.e.* 10^{12} watts) in the energy budget driving this circulation (Munk and Wunsch, 1998; Wunsch, 2000). The oceanic biosphere captures solar energy at a rate of ~ 63 TW (Behrenfeld and Falkowski, 1997)[11]. If only a small percentage of this captured solar

[11]Putting this into perspective, the total human power consumption (electricity and oil) is 15 TW, the

energy makes its way into mechanical energy of swimming, the energy budget can apparently be closed. One terawatt corresponds to an average dissipation rate of 10^{-9} W/kg in the deep oceans, where the buoyancy frequency is typically 10^{-3} s^{-1} or less. Thus, a mean buoyancy length scale for the deep ocean is 1 m or greater. However, in the oceans, small organisms far out number larger animals; organisms of size 1 mm or less are about 10^9 times more abundant than animals of size 1 m or greater[12]. The efficiency of small ($\ell < 1$ mm) abundant organisms in mixing is only 10^{-4}. It is only when one comes to larger, but much less abundant, organisms, such as fish and marine mammals ($\ell > 1$ m), that the mixing efficiency approaches its maximum.

Dissipation is the end product of turbulence. It is also the most readily measured turbulence parameter in the ocean. However, important aspects of turbulence – such as mixing – also depend on the larger scales of turbulent motion (Richardson, 1926; Okubo, 1971).

Measurements of ε in a swarm of krill or school of fish may very well be higher than background. This does not mean this the mixing is proportionally higher as well. Mixing efficiency is a property of the entire turbulent energy spectrum, not only dissipation.

By whatever means one approaches the calculation of biomixing of the oceans, one will always be confronted by the fact that the mixing efficiency of small organisms is extremely low (Visser, 2007a). Most of the mechanical energy they impart to the oceans is dissipated almost immediately as heat. There may be a case to be made for biomixing by larger animals on a local scale, but their relatively low abundance means that they are unlikely to be important contributors to global circulation.

globally averaged geothermal heat supply from the earth's interior is 44 TW, while the poleward heat transport by the atmosphere and oceans is some 4000 TW. These pale in comparison to the solar heat flux the earth receives – about 174,000 TW

[12] This is an estimate drawn from Sheldon's biomass spectrum of oceanic organisms.

5 Motility

The spatial distribution of zooplankton is effected by both their active locomotion and passive drift due to currents, turbulence and circulation cells. Their active locomotion is composed of deterministic elements such as vertical migration, and seemingly random components such as pause-travel, hop-sink, meandering, looping and zigzag trajectories. We have already seen in Sec. 2.2 that random walk motility, either as discrete run-tumble cycles or in a continuous sense, leads to a population description that at one level appears diffusive.

I have presented this primarily as it pertains to encounter rates and for the condition that the parameters that determine the random walk (speed v, correlation length scale λ, and correlation time scale τ) are spatially uniform. Of course, there is no reason to suppose that these parameters do remain uniform. How organisms move – how fast they swim and how often they change direction – are plastic behaviours and can change in response to environmental spatially-varying stimuli. A classic example of this is chemotaxis. Bacteria for instance may swim faster or slower, or turn more often when they encounter an elevated concentration of some chemical. This behaviour causes bacteria to effectively climb chemical gradients, even though they have no sense of the direction of the chemical gradient itself. The dynamics of this as expressed in a continuum, either as a probability function or a concentration field of organisms, is subtly different from a diffusive process proper.

We saw, for instance that a random walk has a "diffusivity"

$$D = \frac{1}{n}\frac{v^2\delta}{1-c} = \frac{1}{n}v^2\tau \tag{5.0.1}$$

That is, that a patch of organisms co-located initially at a point, will disperse outwards with their variance increasing linearly with time. However, for a diffusive process proper, one controlled by Ficks laws, time and spatial properties of the probability density function are described by the diffusive equation:

$$\frac{\partial p}{\partial t} = \nabla \cdot (D\nabla p) \tag{5.0.2}$$

But this is not the same as the expression describing the time and spatial properties of the probability density function of random walking particles. These are described by a similar, but subtlety different Fokker-Planck equation,

$$\frac{\partial p}{\partial t} = \nabla^2 (Dp) \tag{5.0.3}$$

That this does not conform to a diffusive process proper has important consequences for the spatial distribution of organisms. In particular, consider particles with behaviour that varies in response to an environmental stimulus that is distributed un-evenly in space. For a diffusive

process governed by Eq. 5.0.2 the distribution of particles will homogenize eventually having a probability density function that is spatially uniform - irrespective of how their behaviour changes. However, for a random walk governed by Eq. 5.0.3, particles will start to accumulate in areas where the diffusivity is low, eventually attaining a spatial distribution that is inversely proportional to the local diffusivity. This, for instance, is the mechanism often identified as that underlying chemotaxis.

I stress however that this is by no means the whole story. There is a demon deep in the foundations of stochastic calculus – the mathematical formalism underpinning random walks and their associated Fokker-Planck descriptions. Most important for this work: subtle differences in the rules by which random walk behaviour is enacted by individual zooplankters leads to radically different distribution at their population or community level.

The problem of variable random walks is a subject that comes up time and again in different research fields ranging from cosmology, meteorology, economics as well as plankton ecology. Part of the problem is that random walks, at first glance appear to be absolutely transparent. After all, what could be simpler than to give particles (molecules, organisms, stock prices, galaxies) a random kick every now and then and see where they end up. This is the principal underlying Monte Carlo simulations. If the magnitude of the kick varies from place to place, what can that matter? Well matter it does, and to a great extent!! In fact, it is precisely how these random kicks vary from place to place, and how this variation is simulated, that primarily determines the form of the macro-scale properties of the evolving distribution.

Figure 5.1: Monte Carlo simulations – a powerful tool to investigate large dynamic systems – that has been used widely and successful since its inception by Stanislaw Ulam, Ul Enrico Fermi and Jon von Neumann in the 1940's to characterize the properties of the then newly discovered neutrino. Since that time it has been used in numerous fields from macro economics to genetic algorithms. It is named after the casino in Monte Carlo in deference to role played by random processes.

In oceanography for instance, turbulence causes material to disperse in a diffusive manner, and thus may be simulated by a random walk. Furthermore, turbulence varies from place to place. In the vertical for instance, turbulent diffusivity may be high near the surface and bottom where momentum is exchanged across the ocean's boundary; wind stresses at the surface for instance and bed friction at the bottom. In a stratified water column, turbulent diffusivity will tend to be low in the pycnocline where part of the turbulent kinetic energy is dissipated in raising the potential energy of the stratified layer. Within such a vertical profile of diffusivity, particles executing a random walk in a naive sense would accumulate within the bands of low diffusivity (Fig. 5.2). If the particles being simulated were phytoplankton cells, one might be tempted to see this aggregation as a mechanism generating the oft seen subsurface chlorophyll maximum. However, the nature of the particles being simulated should have no bearing on the out come. They could just as likely be salt molecules in which case the aggregation would imply an increase in salinity along the pycnocline. This is patently absurd.

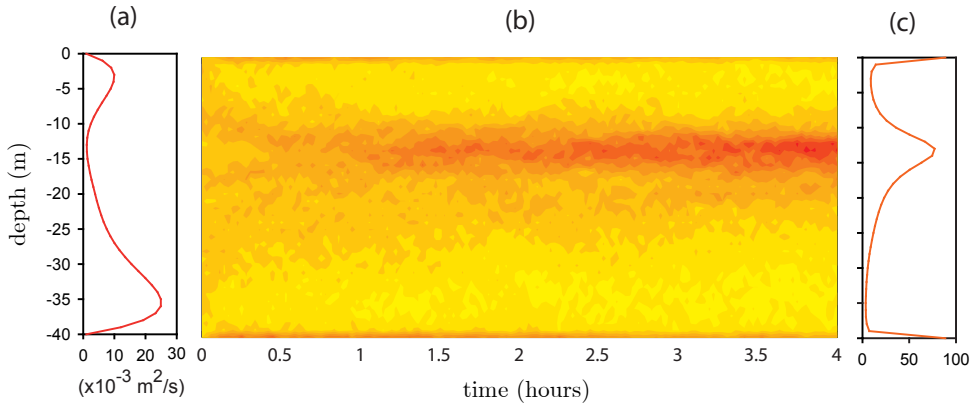


Figure 5.2: A naïve random walk enacted on a initially uniform distribution of particles. The vertical profile of diffusivity D (a) is consistent with turbulent eddy diffusivity in a shallow stratified shelf sea. In the course of 4 hours, particles aggregate (red $\Rightarrow$ high concentration, yellow $\Rightarrow$ low concentration) in the low diffusivity zones (b) at the pycnocline, as well as at the upper and lower boundaries. This behaviour is consistent with the associated Fokker-Planck equation which predicts (c) a steady state particle concentration $C \propto D^{-1}$. This cannot, however, be validated as it violates the well mixed condition.

It was this problem that first drew my attention to random walks, motility, turbulence and diffusivity. The problem I found could be easily solved by applying a "correction" term (Visser, 1997), so that the evolving distribution did not violate the well-mixed condition. This condition, due to Thomson (1987), says that a well-mixed distribution cannot be un-mixed by a random walk simulation. This seems sensible enough – the laws of thermodynamics state that the entropy (*i.e.* disorder) of a system should increase. At the same time however, there has been a long tradition in regarding observed taxis of organisms as arising from a random walk with spatially varying parameters. The well-mixed condition need not apply to organisms as they have their own energy that can be used to decrease entropy. It is this *ad hoc*, *a postiori* "correction" or "non-correction" of random walk simulation to conform to some preconceived macro-scale result that I have found far from satisfactory, and it is in part to address this that I write this chapter.

I have a practical motive as well. Lagrangian simulations and individual based models (IBMs) are becoming increasingly widespread in investigating marine processes, and provide a natural platform upon which the interactions (oftentimes nonlinear) of plankton with each other and their environment can be modelled. The prevalence of affordable and powerful computing capacity has made this approach not only feasible, but readily accessible. Random walks form one of the cornerstones of Lagrangian and individual based models. They are used to simulate diffusive-like processes, and provide a transparent means of modelling both turbulent dispersion (*e.g.* Yamazaki and Kamykowski, 1991; Kamykowski et al., 1994; Franks and Marra, 1994; Visser, 1997) as well as behavioural aspects of swimming (*e.g.* Davis et al., 1991; Turchin, 1998; Flierl et al., 1999; Leising, 2001).

Before random walk models can be employed however, there is the serious consideration of their validation to be considered. As pointed out above, random walks are deceptively simple. They are simple in that they can be written in a few lines of code. But they are also deceptive, in that there are a number of subtle yet crucial implementation and interpretation

issues that have to be considered.

Finally, while random walk models are useful tools in Lagrangian simulations, they also have a strong conceptual appeal. In particular, one of the most powerful aspects of random walks is that they connect the kinematics of individual organisms at the micro-scale with macro-scale dynamics of a population of like organisms. Properties of individual motion can be related directly to a continuum description of concentration fields in advection-diffusion-like equations (more properly Fokker-Planck equations), the currency of traditional Eulerian based models.

5.1 Random walk to diffusion and back again

We have seen that simple random walks, whether discrete or continuous, can be characterized by an equivalent diffusivity

$$D = \frac{1}{n}\frac{v^2\delta}{1-c} \tag{5.1.1}$$

(Sec. 2.2.2). Conversely, a diffusive process can be simulated by a random walk. That is, given a diffusivity D, the trajectory of individual particles, $x(t)$, can be simulated as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{r}_i\sqrt{2nD\delta} \tag{5.1.2}$$

where $\mathbf{x}_i = \mathbf{x}(t = i\delta)$, δ is a short time interval, and $\mathbf{r}_i$ is a random vector in n space such that the expectation value of its moments satisfy $\langle \mathbf{r}_i \rangle = 0$, $\langle \mathbf{r}_i \cdot \mathbf{r}_i \rangle = 1$, and $\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle = 0$ for $i \neq j$. From this, it can be readily shown that the evolution of the variance (2^{nd} moment) of the particle distribution follows

$$\left\langle (\mathbf{x}_m - \mathbf{x}_0)^2 \right\rangle = 2nDt \tag{5.1.3}$$

where $t = m\delta$. That is, variance increases linearly with time – a characteristic of a diffusive process. In fact, Eq. 5.1.3 can serve as a definition of diffusivity D, it is $1/(2n)$ times the rate of change of variance of a particle distribution (Taylor, 1921).

A random walk characterized and realized in terms of diffusivity equates to the Fokker-Planck equation in the continuum. That is, if the position x of a particle executing a random walk is governed in n dimensions by

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{u}(\mathbf{x}_i)\delta + \mathbf{r}_i\sqrt{2nD(\mathbf{x}_i)\delta} \tag{5.1.4}$$

then in the long time limit ($t \gg \delta$), the evolution of the probability density function (or concentration) is described by

$$\frac{\partial p}{\partial t} = -\nabla \cdot (\mathbf{u}p) + \nabla^2(Dp) \tag{5.1.5}$$

For the 3 dimensional case, $p(x, y, z, t)dxdydzdt$ is the probability of finding the particle within a small volume of space $[x \pm dx, y \pm dy, z \pm dz]$ in a small time interval $[t, t + dt]$, although p can also be interpreted as the concentration of a population of like random walkers, provided they do not interact with each other. For generality, an advection velocity $\mathbf{u}$ is included. This result, (*i.e.* Eq. 5.1.4 $\Rightarrow$ Eq. 5.1.5) is well known and its derivation can be found in a number of places (*e.g.* Chapman, 1928; Skellam, 1951; Patlak, 1953; Lapidus and Levandowsky, 1981; Turchin, 1998; Okubo and Levin, 2001).

An important point to note is that Eq. 5.1.5 is not the advection-diffusion equation. It contains an effective additional advective term, $\nabla \cdot (p \nabla D)$, transporting particles down gradients of diffusivity. A consequence of this is that at steady state, $p \propto D^{-1}$. That is, random walking particles following a trajectory defined by Eq. 5.1.4 will aggregate in regions where diffusivity is low. For physical diffusion, (*e.g.* turbulent diffusion) this is deemed non-physical, violating the well-mixed condition (Thomson, 1987; Sawford, 2001) and necessitating a "correction" term (Hunter et al., 1993; Visser, 1997).

By examining the 2^{nd} order moments of the position of random walking particles, it is relatively straight forward to deduce a "correction" to maintain a well-mixed distribution. Specifically, a random walk complying with

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{u}(\mathbf{x}_i) + \nabla D(\mathbf{x}_i))\,\delta + \mathbf{r}_i \sqrt{2nD(\mathbf{x}_i)\delta} \tag{5.1.6}$$

will remain well mixed (Visser, 1997). But what is the rational behind this term, and what is the criterion for its inclusion. To understand this, we have to delve deep into the mathematical formalism of diffusion and random walks, and will uncover some disturbing facets – that certain aspects of random walks and diffusion cannot be derived from first principals. The mathematical basis is simply incomplete.

5.2 Wiener process and stochastic calculus

Wiener processes are continuous-time stochastic (*i.e.* random) processes. They have a very close connection to random walks and are sometimes referred to as Brownian motion – at least in a mathematical sense. Wiener processes together with stochastic calculus provide a mathematical frame work within which random walks can be rigorously expressed. A simple stochastic differential equation can be written

$$d\mathbf{x} = \mathbf{u}dt + \beta d\boldsymbol{\eta} \tag{5.2.1}$$

where u is termed the drift, and β the intensity of the stochastic process. If $\mathbf{x}$ denotes the position of a particle, then the first part of this differential equation is quite straight forward; it states that the rate of change of the position of the particle is its velocity $\mathbf{u}$. The second part represents a stochastic influence on the path of the particle. Specifically, $\boldsymbol{\eta}(t)$ is termed a Wiener process, and has the special properties that

$$\langle d\boldsymbol{\eta} \rangle = 0 \tag{5.2.2}$$

$$\langle d\boldsymbol{\eta} \cdot d\boldsymbol{\eta} \rangle = dt \tag{5.2.3}$$

Given these definitions, it is clear that

$$d\langle \mathbf{x} \rangle = \mathbf{u}dt \tag{5.2.4}$$

$$d\left\langle |\mathbf{x} - \langle \mathbf{x} \rangle|^2 \right\rangle = \beta^2 dt \tag{5.2.5}$$

That is, β^2 is the rate of change of variance of the particle position: a quantity directly related to diffusivity D and number of dimensions n as

$$\beta = \sqrt{2nD} \tag{5.2.6}$$

An excellent treatment of stochastic processes can be found in Crispin Gardiner's book "*Handbook of Stochastic Methods*" (Gardiner, 2004).

5.3 The Chapman-Kolmogorov master equation

Trying to predict the evolution of a probability density function as determined by a set of transition rules may at first seem daunting. There is a pretty neat trick in approaching these sorts of problems, first proposed by Chapman (1928), and again independently by Kolmogorov, which is to look backwards in time a short interval, and deduce how a system came to be in a particular state. That is, if we observe something here and now (by "here" I mean either a physical location or a more general state), it must have arrived according to some well-defined transition rules, from some other locations (states).

5.3.1 Simple 1-D example

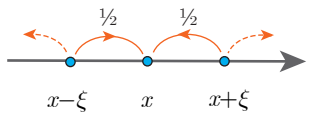


Figure 5.3: A one dimensional random walker moving a distance ξ in a time step δ. If the process is symmetric, one half of the particles arriving at x originate from $x-\xi$ while the other half came from $x+\xi$.

A simple 1-D case illustrates this. Suppose we have random jumpers that can move left or right along a horizontal line (Fig. 5.3). In a short time interval δ they move randomly a distance $\pm\xi$. That is, the probability of finding a particle at x at time $t+\delta$ is

$$p(x,t+\delta) = \frac{1}{2}p(x-\xi,t) + \frac{1}{2}p(x+\xi,t) \tag{5.3.1}$$

where half come from $x-\xi$ and move to the right, and the other half come from $x+\xi$. (The other half leaving these locations move in opposite directions ending up at $x-2\xi$ and $x+2\xi$ and are of no immediate concern here.) If the probability function is continuous in time and space – as it should be – then the terms in this equation can be expanded as a Taylor series

$$p(x,t+\delta) = p(x,t) + \delta\frac{\partial p}{\partial t} + O(\delta^2) \tag{5.3.2}$$

and

$$p(x\pm\xi,t) = p(x,t) \pm \xi\frac{\partial p}{\partial x} + \frac{\xi^2}{2}\frac{\partial^2 p}{\partial x^2} + O(\xi^3) \tag{5.3.3}$$

Substituting these in gives

$$\frac{\partial p}{\partial t} = \frac{\xi^2}{2\delta}\frac{\partial^2 p}{\partial x^2} + O\left(\delta, \frac{\xi^3}{\delta}\right) \tag{5.3.4}$$

As δ becomes small, we can imagine that the higher order terms vanish, and we are left with something that looks very like a diffusion equation. In order for this to be rigorous however, we have to note the following. In the limit as $\delta \to 0$, the variance of the jump distance $\xi^2 \to 0$ also, but it can do so faster or slower than δ. If faster, then $\partial p/\partial t = 0$, if slower then $\partial^2 p/\partial x^2 = 0$, both of which are uninteresting. The nontrivial case is where $\xi^2 \sim \delta$, such that their ratio remains finite as $\delta \to 0$. A random process with this property is known as a Wiener process, and the ratio is the diffusivity:

$$D = \lim_{\delta \longrightarrow 0} \frac{\xi^2}{2\delta} \tag{5.3.5}$$

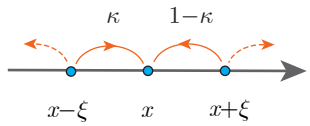


Figure 5.4: An asymmetric random walk where a fraction κ move to the right and the reminder $1-\kappa$ move to the left.

5.3.2 A biased 1-D example

Any number of situations can be analysed in this manner. For instance we could relax the condition that there is an equal probability of jumping left or right and replace it with a bias so that κ jump right while $1-\kappa$ jump left. Thus Eq. (5.3.1) becomes:

$$p(x,t+\delta) = \kappa\, p(x-\xi,t) + (1-\kappa)\, p(x+\xi,t) \tag{5.3.6}$$

which leads to

$$\frac{\partial p}{\partial t} = \frac{(1-2\kappa)\xi}{\delta}\frac{\partial p}{\partial x} + \frac{\xi^2}{2\delta}\frac{\partial^2 p}{\partial x^2} + O\left(\delta, \frac{\xi^3}{\delta}\right) \tag{5.3.7}$$

The factor $(1-2\kappa)$ is the probability of particles going left minus the probability of them going right. If we consider a number of particles co-located at x at time t, then this difference will cause their centre of mass to migrate to the left over time. Specifically, at a time $t+\delta$, their centre of mass is at $x-(1-2\kappa)\xi$. The meaning of $(1-2\kappa)$ thus becomes clear from the general stochastic differential equation, Eq. 5.2.1. Under the assumption that ξ is a Wiener process, the centre of mass of the distribution drifts with a speed u (assumed positive to the right) which gives $u\delta = -(1-2\kappa)\xi$. Thus to order δ^2, the governing Fokker-Planck equation for a "biased" random walk is

$$\frac{\partial p}{\partial t} = u\frac{\partial p}{\partial x} + D\frac{\partial^2 p}{\partial x^2} \tag{5.3.8}$$

and where the bias parameter κ and drift u are related by

$$u = \lim_{\delta\to 0}(1-2\kappa)\frac{\xi}{\delta}$$

5.3.3 The general master equation

Let us consider a particle that can assume a particular state $\mathbf{s}$ composed, for instance of its position $\mathbf{x}$ and velocity $\mathbf{v}$. The state vector may also contain other information. We can then ask the question, what is the probability of finding a particle in state $\mathbf{s}$ at time $t+\Delta t$? If the particle was in state $\mathbf{s}-\Delta\mathbf{s}$ at time t, then this probability can be written as $\psi(\mathbf{s}-\Delta\mathbf{s}, t|\Delta\mathbf{s},\Delta t)$. This is known as the transfer function and is a joint probability that the particle "moves" a distance in state space $\mathbf{s}$ in the time interval Δt given that it was at $\mathbf{s}-\Delta\mathbf{s}$ at time t. The probability of being at $\mathbf{s}-\Delta\mathbf{s}$ at time t is written as $p(\mathbf{s}-\Delta\mathbf{s},t)$. Thus, the probability of the particle being in state $\mathbf{s}$ at time $t+\Delta t$ is the integral over all possible "paths" $\Delta\mathbf{s}$ that lead to $\mathbf{s}$ at time $t+\Delta t$. That is

$$p(\mathbf{s},t+\Delta t) = \int p(\mathbf{s}-\Delta\mathbf{s},t)\psi(\mathbf{s}-\Delta\mathbf{s}|\Delta\mathbf{s},\Delta t)d\Delta\mathbf{s} \tag{5.3.9}$$

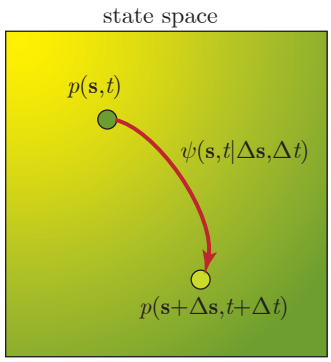


Figure 5.5: The general transfer of particles in state space. The probability of being in state $\mathbf{s}+\Delta\mathbf{s}$ at time $t+\Delta t$ is the integral of the transfer probability $\psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)$ from all possible states $\mathbf{s}$ at time t.

Expanding p and ψ as Taylor series both in time and state space gives

$$\begin{aligned} p+\frac{\partial p}{\partial t}+\cdots & \\ &= \int p\psi d\Delta\mathbf{s} \\ &- \sum_i \int \Delta s_i \frac{\partial}{\partial s_i}\left(p\psi\right) d\Delta\mathbf{s} \\ &+ \frac{1}{2}\sum_i\sum_j \int \Delta s_i \Delta s_j \frac{\partial^2}{\partial s_i s_j}\left(p\psi\right) d\Delta\mathbf{s}+\cdots \end{aligned} \tag{5.3.10}$$

where for brevity, $p \Rightarrow p(\mathbf{s},t)$ and $\psi \Rightarrow \psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)$. Clearly, a particle that was at $\mathbf{s}$ at some time t must still be somewhere in state space at time $t+\Delta t$, Thus

$$\int \psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)d\Delta\mathbf{s}=1 \tag{5.3.11}$$

Integrating the first term in Eq. 5.3.10 we can note that $p(\mathbf{s},t)$ is independent of $\Delta\mathbf{s}$ so that

$$\int p(\mathbf{s},t)\psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)d\Delta\mathbf{s}=p(\mathbf{s},t)\int \psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)d\Delta\mathbf{s}=p(\mathbf{s},t) \tag{5.3.12}$$

Elements of the 2nd term in Eq. 5.3.10 become

$$\begin{aligned} \int \Delta s_i \frac{\partial}{\partial s_i}\left(p(\mathbf{s},t)\psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)\right) d\Delta\mathbf{s} & \\ &= \frac{\partial}{\partial s_i}\left\{p(\mathbf{s},t)\int \Delta s_i \psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)d\Delta\mathbf{s}\right\} \\ &= \frac{\partial}{\partial s_i}\left\{p(\mathbf{s},t)\langle\Delta s_i\rangle\right\} \end{aligned} \tag{5.3.13}$$

where we have noted that the expectation value of Δs_i is the integral of its product with its transfer function over all possible values. Finally, by similar arguments, elements of the 3rd term in Eq. 5.3.10 are

$$\begin{aligned} \int \Delta s_i \Delta s_j \frac{\partial^2}{\partial s_i s_j}\left(p(\mathbf{s},t)\psi(\mathbf{s},t|\Delta\mathbf{s},\Delta t)\right) d\Delta\mathbf{s} & \\ &= \frac{\partial^2}{\partial s_i s_j}\left\{p(\mathbf{s},t)\langle\Delta s_i \Delta s_j\rangle\right\} \end{aligned} \tag{5.3.14}$$

Thus, the general Fokker-Planck equation in phase space can be written

$$\frac{\partial p}{\partial t} = -\sum_i \frac{\partial}{\partial s_i}\{pA_i\} + \sum_i \sum_j \frac{\partial^2}{\partial s_i s_j}\{pD_{ij}\} \tag{5.3.15}$$

where the general drift term is

$$A_i = \lim_{\Delta t \to 0} \frac{\langle \Delta s_i \rangle}{\Delta t} \tag{5.3.16}$$

and the general diffusion is

$$D_{ij} = \lim_{\Delta t \to 0} \frac{\langle \Delta s_i \Delta s_j \rangle}{\Delta t} \tag{5.3.17}$$

This is sometimes referred to as the backward Kolmogorov equation.

The application of most interest here is that where the state vector $\mathbf{s}$ is dependent only on position $\mathbf{x}$, and that position is driven by a Wiener process $\boldsymbol{\eta}$. That is

$$\Delta \mathbf{x} = \mathbf{u}\Delta t + \beta \Delta \boldsymbol{\eta} \tag{5.3.18}$$

Thus

$$\langle \Delta x_i \rangle = u_i \Delta t \tag{5.3.19}$$

and

$$\langle \Delta x_i \Delta x_j \rangle = \begin{cases} \beta^2 \Delta t & i = j \\ 0 & i \neq j \end{cases} \tag{5.3.20}$$

so that the associated Fokker-Planck equation can be written

$$\frac{\partial p}{\partial t} = -\nabla \cdot (\mathbf{u}p) + \nabla^2 (Dp) \tag{5.3.21}$$

5.3.4 The telegraph equation

We are now in a position to derive the telegraph equation, introduced in Sec. 2.3.4 from first principals. Consider a number of particles that move either left or right along the x axis. In a short time interval δ, the distance they move is $\xi = v\delta$ where v is speed of the particles, assumed to be constant (Fig. 5.6). In the same time interval, the particles have a probability of changing direction $\gamma = \alpha\delta$, and thus a probability of not changing direction $1 - \alpha\delta$. We can split the particles into two sub-populations: those moving to the right with concentration $p^+(x,t)$ and those moving to the left with concentration $p^-(x,t)$. The total concentration of particles is thus $p(x,t)$. The question as to how the concentration (or equivalently probability density function) evolves over time, can be deduced from the Chapman-Kolmogorov master equation applied to each sub-population.

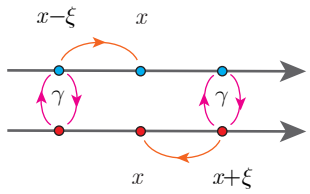


Figure 5.6: Derivation of the telegraph equation can be got by applying the Chapman-Kolmogorov master equation to two populations of particles, one moving right and the other moving left with probability of changing from one to the other of γ.

5.3.5 Simple 1-D example

Specifically, a right moving particle seen at a given location x at a time $t+\delta$ must have come from the particles located at $x-\xi$ at time t, either as particles that were moving right and continued to do so, or as left moving particles that changed direction. That is,

$$p^+(x,t+\delta) = (1-\gamma)p^+(x-\xi,t) + \gamma p^-(x-\xi,t) \tag{5.3.22}$$

Similarly, the master equation for left moving particles can be written

$$p^-(x,t+\delta) = (1-\gamma)p^-(x+\xi,t) + \gamma p^+(x+\xi,t) \tag{5.3.23}$$

Expanding these as Taylor series up to explicit $O(\delta^2)$ and $O(\xi^2)$ gives

$$\delta\frac{\partial p^+}{\partial t} = -\xi\frac{\partial p^+}{\partial x} + \xi\gamma\frac{\partial}{\partial x}\left(p^+ - p^-\right) - \gamma\left(p^+ - p^-\right) \tag{5.3.24}$$

$$\delta\frac{\partial p^-}{\partial t} = -\xi\frac{\partial p^-}{\partial x} + \xi\gamma\frac{\partial}{\partial x}\left(p^+ - p^-\right) + \gamma\left(p^+ - p^-\right) \tag{5.3.25}$$

It should be noted that both xi and γ are order δ so that in the limit as $\delta \to 0$, the second term in both equations disappear, and the resulting differential equations become

$$\frac{\partial p^+}{\partial t} = -v\frac{\partial p^+}{\partial x} - \alpha\left(p^+ - p^-\right) \tag{5.3.26}$$

$$\frac{\partial p^-}{\partial t} = v\frac{\partial p^-}{\partial x} + \alpha\left(p^+ - p^-\right) \tag{5.3.27}$$

By taking the sum and difference of these equations, noting $p = p^+ + p^-$ and defining $q = p^+ - p^-$, these give

$$\frac{\partial p}{\partial t} = -v\frac{\partial q}{\partial x} \tag{5.3.28}$$

$$\frac{\partial q}{\partial t} = -v\frac{\partial p}{\partial x} - 2\alpha q \tag{5.3.29}$$

Finally, taking $\partial/\partial t$ of the first and $\partial/\partial x$ and manipulating gives

$$\frac{\partial^2 p}{\partial t^2} + 2\alpha\frac{\partial p}{\partial t} = v^2\frac{\partial^2 p}{\partial x^2} \tag{5.3.30}$$

which is the telegraph equation.

5.3.6 Variable pause time

We now turn to the class of random walks where their governing parameters vary in space. Consider the case where particles can move either left or right or pause(Fig. 5.7). Furthermore let us suppose that the probability of pausing is variable in space.

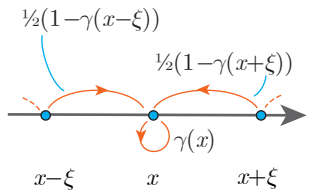


Figure 5.7: The probability of a random walker not moving is $\gamma(x)$ and is in general a function of space.

The Chapman-Kolmogorov master equation for this process can be written

$$p(x,t+\delta) = \gamma p(x,t) + \frac{1}{2}(1-\gamma(x-\xi))p(x-\xi,t) + \frac{1}{2}(1-\gamma(x+\xi))p(x+\xi,t) \tag{5.3.31}$$

Expanding both p and γ as Taylor series gives

$$p + \delta\frac{\partial p}{\partial t} = \gamma p + \frac{1}{2}\left(1-\gamma+\xi\frac{\partial\gamma}{\partial x} - \frac{1}{2}\xi^2\frac{\partial^2\gamma}{\partial x^2}\right)\left(p - \xi\frac{\partial p}{\partial x} + \frac{1}{2}\xi^2\frac{\partial^2 p}{\partial x^2}\right) + \frac{1}{2}\left(1-\gamma-\xi\frac{\partial\gamma}{\partial x} - \frac{1}{2}\xi^2\frac{\partial^2\gamma}{\partial x^2}\right)\left(p + \xi\frac{\partial p}{\partial x} + \frac{1}{2}\xi^2\frac{\partial^2 p}{\partial x^2}\right) \tag{5.3.32}$$

which can be reduced to

$$\delta\frac{\partial p}{\partial t} = \frac{1}{2}\xi^2\frac{\partial^2}{\partial x^2}\left((1-\gamma)p\right) \tag{5.3.33}$$

The function $1-\gamma$ is a measure of the percentage of particles that are actively moving. Thus the local diffusivity depends only on this proportion, the remaining faction γ does not move. That is

$$D = \lim_{\delta\to 0}\frac{1}{2}\xi^2(1-\gamma) \tag{5.3.34}$$

so that the governing Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = \frac{\partial^2}{\partial x^2}\left(Dp\right) \tag{5.3.35}$$

5.3.7 Variable run length

Consider a one dimensional random walker (left – right – no pause) that changes the length of its run as a function of space (Fig. 5.8). The Chapman-Kolmogorov master equation corresponding to this is

$$p(x,t+\delta)dx = \frac{1}{2}p(x-\xi^-,t)dx^- + \frac{1}{2}p(x+\xi^+,t)dx^+ \tag{5.3.36}$$

where the run lengths $\xi^\pm$ are governed by a Wiener process:

$$\xi^- = \beta(x-\xi^-)|\dot\eta| \tag{5.3.37}$$
$$\xi^+ = \beta(x+\xi^+)|\dot\eta| \tag{5.3.38}$$

Here for brevity, I use the notation $\dot\eta = \eta(t+\delta) - \eta(t)$.

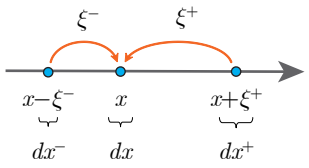


Figure 5.8: The run length of a random walker $\xi(x)$, is in general a function of space. That is the intensity β, or equivalently the diffusivity D can vary in space. Here, particles arriving at x from the right travel further ξ^+ than those arriving from the left ξ^-. Because of the anisotropy of travel distance, the probability density function, *i.e.* probable number per unit distance, must take into account variable departure and arrival intervals dx, dx^- and dx^+.

I have explicitly written the spatial interval dx in the master equation. Importantly, because the run lengths vary is space, the interval from which particles arriving at $x \pm dx$ come from may be greater or smaller than $2dx$; hence the distinction between dx, dx^- and dx^+ in Eq. 5.3.36. A Taylor expansion of ξ^- and ξ^+ gives

$$\xi^- = \beta|\dot{\eta}| - \frac{\partial D}{\partial x}\delta \tag{5.3.39}$$

$$\xi^+ = \beta|\dot{\eta}| + \frac{\partial D}{\partial x}\delta \tag{5.3.40}$$

from which we can derive

$$dx^- = d(x - \xi^-) = dx\left(1 - \frac{\partial \beta}{\partial x}\beta|\dot{\eta}| + \frac{\partial^2 D}{\partial x^2}\delta\right) \tag{5.3.41}$$

$$dx^+ = d(x + \xi^+) = dx\left(1 + \frac{\partial \beta}{\partial x}\beta|\dot{\eta}| + \frac{\partial^2 D}{\partial x^2}\delta\right) \tag{5.3.42}$$

Thus, to $O(\dot{\eta}^3)$,

$$\begin{aligned}\delta\frac{\partial p}{\partial t} &= p\frac{\partial^2 D}{\partial x^2}\delta \\ &+ \frac{\partial p}{\partial x}\left(\frac{\xi^+ - \xi^-}{2} + \frac{\xi^+ + \xi^-}{2}\frac{\partial \beta}{\partial x}|\dot{\eta}|\right) \\ &+ \frac{(\xi^+)^2 + (\xi^-)^2}{4}\frac{\partial^2 p}{\partial x^2}\end{aligned} \tag{5.3.43}$$

This can be simplified by noting the following identities derived from Eq. 5.3.39 and Eq. 5.3.40:

$$\frac{\xi^+ - \xi^-}{2} = \frac{\partial D}{\partial x}\delta \tag{5.3.44}$$

$$\frac{\xi^+ + \xi^-}{2}\frac{\partial \beta}{\partial x}|\dot{\eta}| = \beta\frac{\partial \beta}{\partial x}\delta = \frac{\partial D}{\partial x}\delta \tag{5.3.45}$$

$$\frac{(\xi^+)^2 + (\xi^-)^2}{4} = \frac{1}{2}\beta^2\delta = D\delta \tag{5.3.46}$$

leading to

$$\frac{\partial p}{\partial t} = p\frac{\partial^2 D}{\partial x^2} + 2\frac{\partial D}{\partial x}\frac{\partial p}{\partial x} + D\frac{\partial^2 p}{\partial x^2} \tag{5.3.47}$$

Simplifying even further, this becomes the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{\partial^2}{\partial x^2}(Dp) \tag{5.3.48}$$

Not only the intensity but also the drift of a stochastic process can vary in space. This can be readily included in the previous argument where additional terms $\pm u\delta$ are included in the expressions for $\xi^{\pm}$.

$$\xi^{-} = \beta|\dot{\eta}| - \frac{\partial D}{\partial x}\delta + u\delta \tag{5.3.49}$$

$$\xi^{+} = \beta|\dot{\eta}| + \frac{\partial D}{\partial x}\delta - u\delta \tag{5.3.50}$$

These carry into the expressions for $dx^{\pm}$ with $-\partial u/\partial x\delta$ appearing in both. Carrying these through, the associated Fokker-Planck equation becomes

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(up) + \frac{\partial^2}{\partial x^2}(Dp) \tag{5.3.51}$$

5.4 Itô, Stratonovich and Transport

In stochastic integration, the two variants most often encountered are the Itô integral and the Stratonovich integral. Both are akin to the Riemann-Stieltjes integral in that they are the limit of Riemann sums. For instance, the stochastic differential equation

$$dx = \beta(x)d\eta \tag{5.4.1}$$

implies the existence of an integral

$$x(T) - x(0) = \int_0^T \beta(x(t))d\eta(t). \tag{5.4.2}$$

The interpretation of this in terms of a Riemann-Stieltjes integral is

$$x(T) - x(0) = \lim_{\delta\to 0}\sum_{i=0}^{T/\delta}\beta(x(t_i^*))(\eta(t_i+\delta) - \eta(t_i)) \tag{5.4.3}$$

where

$$t_i^* = t_i + \alpha\delta \tag{5.4.4}$$

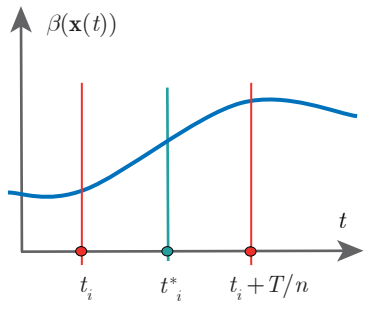


Figure 5.9: Integrating $dx = \beta(x)d\eta$ in terms of a Riemann-Stieltjes integral. The time domain $[0, T]$ can be divided into $n = T/\delta$ time steps of mean duration δ. Thus, $d\eta \sim \eta(t_i + \delta) - \eta(t_i)$ and the integrand can be evaluated anywhere in the interval $[t_i, t_i + \delta]$.

That is, the integrator is a Wiener process, and is well defined by evaluations of η at t_i and $t_i + \delta$. The integrand, on the other hand, is evaluated at a time t_i^* somewhere in the interval

$[t_i, t_i + \delta]$ (Fig. 5.9). It is the time within this interval at which the integrand is evaluated that marks the distinction between Itô and Stratonovich calculus. Specifically Itô integration proceeds with $\alpha = 0$ where the integrand is evaluated at the start of each interval, whereas Stratonovich has $\alpha = 1/2$ where evaluation is at the centre of each interval. As we will see, the choice of α, whether 0, 1/2 or any other value between 0 and 1, has enormous repercussion for the value of the integral. This is the demon at the foundation of stochastic integration.

5.4.1 Fokker-Planck and α-integration

Consider a 1 dimensional stochastic trajectory where particle positions, $x(t)$, evolve according to the stochastic differential equation

$$dx = udt + \beta(x)d\eta \tag{5.4.5}$$

where u and β are respectively the drift and intensity of the stochastic process. It can be shown that the continuum equation describing the time evolution of the probability density function for particle following this general random walk process can be written as

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(\mathbf{u}p) + \frac{\partial^2}{\partial x^2}(Dp) \tag{5.4.6}$$

This is quite general, and takes into account the possibility that u and/or D (and thus β) can vary spatially.

The drift term, u, in the above can be any function. A particularly revealing case is when this term is set to

$$u = \alpha\frac{\partial D}{\partial x} \tag{5.4.7}$$

and where α is a factor between 0 and 1. In this case, the corresponding Fokker-Planck equation becomes

$$\frac{\partial p}{\partial t} = (1-\alpha)\frac{\partial}{\partial x}\left(p\frac{\partial D}{\partial x}\right) + \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) \tag{5.4.8}$$

Thus, when $\alpha = 0$, we have

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(p\frac{\partial D}{\partial x}\right) + \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) = \frac{\partial^2}{\partial x^2}\left(Dp\right) \tag{5.4.9}$$

which leads to aggregation in regions of low D, whereas when $\alpha = 1$, we have

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) \tag{5.4.10}$$

which is pure diffusion, and leads to no aggregation.

The interpretation of α can be got by examining the relationship

$$x(t+\delta) - x(t) = \xi = \beta(x + \alpha\xi)\Delta\eta \tag{5.4.11}$$

This is simply the 1 dimensional discrete interval version of Eq. 5.2.1 where 1) the explicit drift term is dropped and 2) the intensity of the stochastic process is evaluated not at the

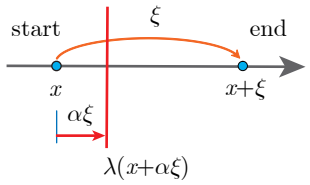


Figure 5.10: The evaluation of the intensity of a stochastic process with the jump interval.

start of each excursion, but at a point between the start and end locations. This is sketched in Fig. 5.10. A Taylor expansion about x thus leads to

$$\begin{aligned} x(t+\delta) - x(t) = \xi &= \beta(x)\Delta\eta + \alpha\xi\frac{\partial\beta}{\partial x}\Delta\eta \\ &= \beta(x)\Delta\eta + \alpha\beta\frac{\partial\beta}{\partial x}\delta \\ &= \beta(x)\Delta\eta + \alpha\frac{\partial D}{\partial x}\delta \end{aligned} \tag{5.4.12}$$

Thus evaluating the intensity of the stochastic process at a position$x + \alpha\xi$ with no drift is equivalent to evaluating the process at the start position x and adding a drift velocity $u = \alpha\partial D/\partial x$. The corresponding Fokker-Planck equation leads to pure diffusion when $\alpha = 1$; when the intensity of the stochastic process is evaluated at the end of each excursion.

5.4.2 Variable run length and α-integration

We can proceed as in Sec. 5.3.7, this time evaluating the displacements not at the point of departure, but at a location somewhere between the start point and the end point (Fig. 5.11). If we take α as the ratio of the position of evaluation to the total displacement, then we can write

$$\xi^- = \beta(x - (1-\alpha)\xi^-)|\dot{\eta}| \tag{5.4.13}$$

$$\xi^+ = \beta(x + (1-\alpha)\xi^+)|\dot{\eta}| \tag{5.4.14}$$

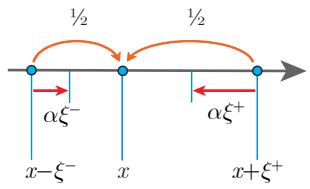


Figure 5.11: Chapman-Kolmogorov arrivals for variable run lengths and where evaluation of the stochastic process is within the interval in accordance with α-integration.

which can be expanded as

$$\xi^- = \beta|\dot{\eta}| - (1-\alpha)\frac{\partial D}{\partial x}\delta \tag{5.4.15}$$

$$\xi^+ = \beta|\dot{\eta}| + (1-\alpha)\frac{\partial D}{\partial x}\delta \tag{5.4.16}$$

and lead to

$$dx^- = d(x-\xi^-) = dx\left(1 - \frac{\partial\beta}{\partial x}\beta|\dot{\eta}| + (1-\alpha)\frac{\partial^2 D}{\partial x^2}\delta\right) \tag{5.4.17}$$

$$dx^+ = d(x+\xi^+) = dx\left(1 + \frac{\partial\beta}{\partial x}\beta|\dot{\eta}| + (1-\alpha)\frac{\partial^2 D}{\partial x^2}\delta\right) \tag{5.4.18}$$

The corresponding equation governing the evolution of the probability density function is then given by

$$\begin{aligned}\delta\frac{\partial p}{\partial t} &= p(1-\alpha)\frac{\partial^2 D}{\partial x^2}\delta \\ &+ \frac{\partial p}{\partial x}\left(\frac{\xi^+ - \xi^-}{2} + \frac{\xi^+ + \xi^-}{2}\frac{\partial \beta}{\partial x}|\dot{\eta}|\right) \\ &+ \frac{(\xi^+)^2 + (\xi^-)^2}{4}\frac{\partial^2 p}{\partial x^2}\end{aligned} \tag{5.4.19}$$

This can be simplified by noting the following identities derived from Eq. 5.4.15 and Eq. 5.4.16:

$$\frac{\xi^+ - \xi^-}{2} = (1-\alpha)\frac{\partial D}{\partial x}\delta \tag{5.4.20}$$

$$\frac{\xi^+ + \xi^-}{2}\frac{\partial \beta}{\partial x}|\dot{\eta}| = \frac{\partial D}{\partial x}\delta \tag{5.4.21}$$

$$\frac{(\xi^+)^2 + (\xi^-)^2}{4} = D\delta \tag{5.4.22}$$

The corresponding equation governing the evolution of the probability density function is thus given by

$$\frac{\partial p}{\partial t} = (1-\alpha)p\frac{\partial^2 D}{\partial x^2} + (2-\alpha)\frac{\partial D}{\partial x}\frac{\partial p}{\partial x} + D\frac{\partial^2 p}{\partial x^2} \tag{5.4.23}$$

Simplifying even further, this becomes the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) + (1-\alpha)\frac{\partial}{\partial x}\left(p\frac{\partial D}{\partial x}\right) \tag{5.4.24}$$

There is something quite remarkable about these results; namely that the evaluation of the stochastic integral is strongly dependent on where the integrand is evaluated within a small interval - even in the limit as the interval tends to zero!!. This is a phenomenon not encountered in ordinary integration and arises because of the nature of the stochastic process: a small change in interval δ corresponds to a change in integrand proportional to $\sqrt{\delta}$ becoming comparatively more important as $\delta \to 0$. From a practical point of view, this means that whenever a stochastic integration is preformed (*e.g.* a random walk implementation) the axiomatic assumption of the integration scheme, whether Itô, Stratonovich or any one of the α-integration scheme continuum, pre-determines the nature of the solution. There is nothing in the mathematical formalism that promotes one specific α-integration scheme over another (Visser, 2008).

In this interpretation, the correction term introduced to maintain the well-mixed condition in random walk simulations of physical diffusion (Hunter et al., 1993; Visser, 1997) is simply transforming from the Itô stochastic integral ($\alpha = 0$) to the Transport integral ($\alpha = 1$); that is, transforming to a stochastic integration scheme that is consistent with physical diffusion. Mathematical rigour cannot be used to distinguish one form the other. Model validation can only come from additional information such as the macroscopic distribution of particles (*e.g.* is the well-mixed condition met?) or a microscopic analysis of organisms behaviour (*e.g.* does translocation exhibit a directional bias?).

5.4.3 Variable pause and α-integration

One might think then from these two sections that the α-integration of a stochastic process is a well founded path to a Fokker-Planck continuum description. Apparently not. In section (5.3.6) we saw that the governing Fokker-Planck equation for the Itô integral ($\alpha = 0$) for the variable pause has exactly the same form as for the Itô integral for the variable run-length. So far so good. But now consider evaluating the pause probability at the position $\alpha\xi$ with each excursion. The master equation can be written

$$\begin{aligned} p(x, t+\delta) = \gamma p(x,t) &+ \frac{1}{2}(1-\gamma(x-(1-\alpha)\xi))p(x-\xi,t) \\ &+ \frac{1}{2}(1-\gamma(x+(1-\alpha)\xi))p(x+\xi,t) \end{aligned} \tag{5.4.25}$$

Expanding both p and γ as Taylor series gives

$$\begin{aligned} p + \delta\frac{\partial p}{\partial t} = &\gamma p \\ &+ \frac{1}{2}\left(1-\gamma+(1-\alpha)\xi\frac{\partial\gamma}{\partial x} - \frac{1}{2}(1-\alpha)^2\xi^2\frac{\partial^2\gamma}{\partial x^2}\right)\left(p - \xi\frac{\partial p}{\partial x} + \frac{1}{2}\xi^2\frac{\partial^2 p}{\partial x^2}\right) \\ &+ \frac{1}{2}\left(1-\gamma-(1-\alpha)\xi\frac{\partial\gamma}{\partial x} - \frac{1}{2}(1-\alpha)^2\xi^2\frac{\partial^2\gamma}{\partial x^2}\right)\left(p + \xi\frac{\partial p}{\partial x} + \frac{1}{2}\xi^2\frac{\partial^2 p}{\partial x^2}\right) \end{aligned} \tag{5.4.26}$$

which can be reduced to

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left((1-2\alpha)p\frac{\partial D}{\partial x}\right) + \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) \tag{5.4.27}$$

That is, the diffusion equation is returned for $\alpha = 1/2$ (the Stratonovich integral) as compared to $\alpha = 1$ (the Transport integral) for the variable run-length case considered above. The rules that govern random walks, and the macros-scale distribution patterns they create are subtle indeed.

5.5 Diffusivity and motility decomposed

Reading up to this point, one might expect that the diffusivity D of a random walk (or equivalently the intensity β of the associated stochastic process) is a well defined parameter whose variation in space has a given solution for a given α-integration. However, D (and it associated β) is a compound parameter. It depends on two micro-scale variables; speed v and correlation time scale τ (or correlation length scale λ). The same spatial variation of diffusivity can be achieved by varying either speed, or timescale, or both. Despite seemingly flawless arguments outlined above, these do not yield the same macroscopic, Fokker-Planck description (Schnitzer, 1993; Visser and Thygesen, 2003).

To get some indication of this difference, we can simulate the random walk not via Eq. 5.1.4 or Eq. 5.1.6, but at a more fundamental level as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + v(\mathbf{x}_i)\hat{\boldsymbol{\phi}}_{i+1}\delta \tag{5.5.1}$$

$$\hat{\boldsymbol{\phi}}_{i+1} = \hat{\boldsymbol{\phi}}_i + \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix} \tag{5.5.2}$$

At each time step of duration δ, an angle θ_i is selected from a normal distribution with mean $\langle\theta_i(\mathbf{x})\rangle = 0$ and variance $\langle\theta_i(\mathbf{x})^2\rangle = \sigma(\mathbf{x})$ where in general, the variance is a function of

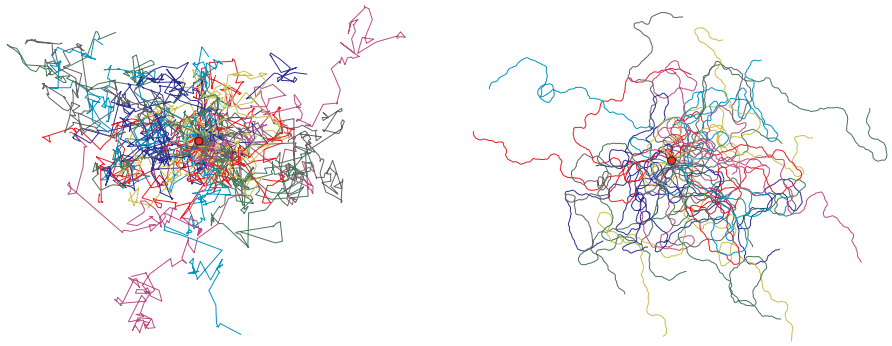


Figure 5.12: Trajectories of 30 particles executing discrete and continuous random walks for the same motility parameters.

position. For a normal distribution, the probability of angles laying in the interval $[\theta, \theta + d\theta]$ is

$$P(\theta, \theta + d\theta) = \frac{d\theta}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\theta^2}{2\sigma}\right) \tag{5.5.3}$$

and $\langle \cos\theta \rangle = e^{-\sigma/2} = c$. Thus, the correlation time scale τ can be related to the variance of the distribution of relative turn angles at time lag δ as

$$\frac{\exp(-\sigma(\mathbf{x}))}{\sigma(\mathbf{x})} = \left(1 - \frac{\delta}{\tau(\mathbf{x})}\right)^2 \tag{5.5.4}$$

In regions where the correlation time scale is large, the angular variance is small; particles tend to travel in long straight paths. Conversely, where this time scale is small, the angular variance is large; particles tend to travel in convoluted paths. Within the above formulation, the correlation time scale (klinokinesis), and the speed of swimming (orthokinesis) can be adjusted independently, in response to an environmental cue.

While this problem can be approached analytically, it is extremely useful to model such processes numerically. Such models can be written in a couple of lines of code, run quickly, and can be used to verify the macroscale effects of particular random walk implementations effects that may be hidden in rigorous but impenetrable mathematical analysis.

In Fig. 5.13 the results of such a model are presented. This represents, for instance, a vertical slice through a thin layer, wherein motile organisms change their local diffusivity by changing either (a) their correlation time scale (klinokinesis), or (b) their swimming speed (orthokinesis). 10,000 particles execute a vertical random walk with speed $v(z)$ and correlation time scale $\tau(z)$. The simulation covers a 2 m vertical section of the water column with a 20 cm thin layer wherein organisms (*i.e.* particles) change their turning frequency or swimming speed. For (a), $\tau(z) = 15$ s outside and 1.5 s inside the thin layer while $v(z) = 1$ cm/s throughout the whole simulated region. For (b), $v(z) = 1$ cm/s outside and 0.316 cm/s inside the thin layer, while $\tau(z) = 15$ s throughout the whole simulation. In both cases, diffusivity has the same variation across the thin layer. After 1 hour simulation time, the klinokinesis case (a) is still homogeneous while the orthokinesis case (b) displays a significant increase in concentration within the thin layer. It is apparent is that the effect of variable swimming speed leads to aggregation while variable turn frequency does not even though

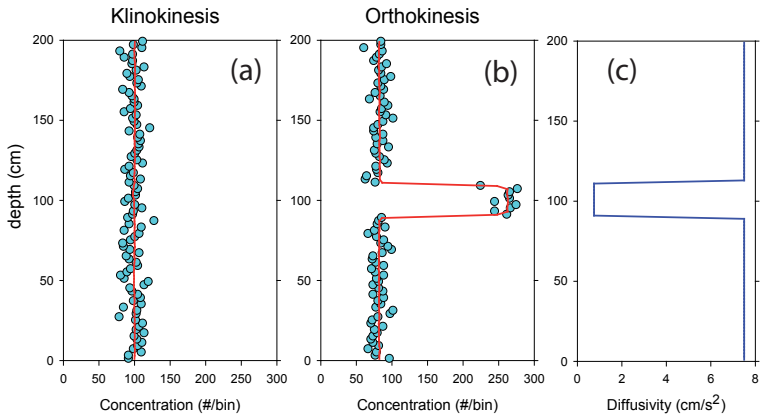


Figure 5.13: Numerical experiment results. Spatially variable random walk where diffusivity (c) has the same profile, but where variable correlation time scale (a: klinokinesis) and translational speed (b: orthokinesis) are examined separately. 10,000 particles are tracked at 1 s time steps. Reflecting boundary conditions are applied at the surface and bottom while periodic boundaries are applied at the sides. The frequency distribution is calculated in 2 cm bins; the symbols giving the distribution at the last time step (after 1 h simulated time), and the solid line giving mean distribution per bin over the previous 1/2 h.

their contribution to variable diffusivity, $D(z) = v(z)^2\tau(z)/n$ is equivalent. That is, spatial variation of in speed and turning frequency do *not* have the same effect on aggregation even when their contribution to spatial variation in diffusivity is identical. The following outlines theoretical reasoning to support this.

5.5.1 Master equation – variable speed and correlation time scale.

Consider motile behaviour which can be divided into different segments: runtumble cycles, although they could equally well be swimpause cycles or jumpsit cycles. Assume that each of these segments is independent of and of identical statistics. In the context of the motility of plankters, this means that the environment is homogenous. The motility D can be seen as an equivalent diffusivity and computed from the statistics of the runtumble cycle as in the central limit theorem. Specifically,

$$D = \frac{\langle \Delta \ell^2 \rangle}{2n \langle \Delta t_c \rangle} \tag{5.5.5}$$

where n is the number of dimensions, $\langle \Delta \ell^2 \rangle$ is the mean square displacement during a runtumble cycle, and $\langle \Delta t_c \rangle$ is the mean duration of a cycle.

For a run-tumble motion, the duration of the run itself is Δt while the duration of the tumble is Δs. The total cycle duration is $\langle \Delta t_c \rangle = \tau + \sigma$ where $\tau = \langle \Delta t \rangle$ and $\sigma = \langle \Delta s \rangle$. If furthermore the speed v is constant during the run, then $\langle \Delta \ell^2 \rangle = v^2 \langle \Delta t^2 \rangle$. Combining terms, we obtain

$$D = \frac{v^2}{2n} \frac{\langle \Delta t^2 \rangle}{\langle \Delta t \rangle + \langle \Delta s \rangle} \tag{5.5.6}$$

Assume that the run lengths are exponentially distributed. In this case $\langle \Delta t^2 \rangle = 2\tau^2$ so

$$D = \frac{v^2}{n} \frac{\tau}{1 + \sigma/\tau} \tag{5.5.7}$$

However, in general the distribution of run durations may be over-dispersed or under-dispersed relative to the exponential distribution. In this context it is convenient to restate Eq. 5.5.6 as

$$D = \frac{v^2\tau}{2n}\frac{1+CV^2}{1+\sigma/\tau} \tag{5.5.8}$$

where CV is the coefficient of variation of run lengths, *i.e.* the ratio of the standard variation to the mean. As an extreme example, all run lengths may have identical length, *i.e.* $CV = 0$, which halves the diffusivity compared with exponential run lengths. At the other extreme, the mean run length $\tau = \langle \Delta t \rangle$ may be finite but the mean squared run length $\langle \Delta \ell^2 \rangle$ may be infinite, *e.g.* if the tail of the probability density behaves like $1/t^p$ for $2 < p \leq 3$ (Levy-like statistics, *cf.* Sec. 6.4). In this case no equivalent diffusivity exists (Metzler and Klafter, 2000).

Now assume that the motion takes place in a heterogeneous environment, *i.e.* the parameters in Eq. 5.5.5 vary in space. Strict compliance with the central limit theorem requires a separation of time scales. Specifically, the spatial heterogeneity of random walk parameters must be small in the course of a single run, so that an intermediate time scale exists in which the mean-square displacement grows linearly with time. This time scale is large compared to the duration of a single run-tumble cycle, and yet changes in the parameters are small over this time scale. The local diffusivity at each point can thus be estimated by inserting the local values of the parameters. For simplicity, consider first the case of no tumble time $\langle \Delta s \rangle = 0$ and exponentially distributed run lengths so that the coefficient of variation is one. Then the local diffusivity simplifies to $D(\mathbf{x}) = v(\mathbf{x})^2\tau(\mathbf{x})/n$. The state of the individual plankter at time t is then fully described by its position $\mathbf{x}^\bullet$ and the direction $\boldsymbol{\phi}^\bullet$ with which it moves:

$$d\mathbf{x}(t) = v(\mathbf{x}(t))\boldsymbol{\phi}(\mathbf{x}(t))dt \tag{5.5.9}$$

where the direction $\boldsymbol{\phi}^\bullet$ is constant during a run. Here, Lagrangian variables are denoted by a superscript $^\bullet$, *i.e.* following an organism.

Let $\psi(\mathbf{x}, \boldsymbol{\phi}, t)$ be the density in state space, *i.e.* $\psi(\mathbf{x}, \boldsymbol{\phi}, t)d\mathbf{x}d\boldsymbol{\phi}$ is the number of individuals which at time t that have position $\mathbf{x}$ and direction $\boldsymbol{\phi}$. Then ψ satisfies the Master equation (Gardiner, 2004)

$$\frac{\partial \psi}{\partial t} = -\nabla \cdot (v\boldsymbol{\phi}\psi) - \alpha\psi + \int \alpha\psi(\mathbf{x}, \boldsymbol{\phi}', t)d\boldsymbol{\phi}' \tag{5.5.10}$$

where $\alpha(\mathbf{x}) = 1/\tau(\mathbf{x})$ is the tumble rate. The master equation expresses the balance of numbers in an infinitesimal cell in phase space. The term $v\boldsymbol{\phi}\psi$ appears as an advective flux in state space and its divergence gives rise to local depletion. The term $\alpha\psi$ represents a local sink as particles tumble out of the state $(\mathbf{x} \pm d\mathbf{x}, \boldsymbol{\phi} \pm d\boldsymbol{\phi})$. Likewise, the 3rd term represents the local source of particles that tumble into the state $(\mathbf{x} \pm d\mathbf{x}, \boldsymbol{\phi} \pm d\boldsymbol{\phi})$. Note that the integral over directions is normalized so that $\int d\boldsymbol{\phi} = 1$.

The steady-state solution to the Master equation is

$$\psi(\mathbf{x}, t) = \frac{\psi_0}{v(\mathbf{x})} \tag{5.5.11}$$

This can be verified by inserting into the Master equation: the advective flux becomes $\boldsymbol{\phi}\psi$, which is independent of $\mathbf{x}$, so the divergence term in Eq. 5.5.10 vanishes. As for the integral

in Eq. 5.5.10, the integrand is constant w.r.t. $\boldsymbol{\phi}'$, hence

$$-\alpha\psi + \int \alpha\psi d\boldsymbol{\phi}' = -\alpha\frac{\psi_0}{v(\mathbf{x})} + \alpha\frac{\psi_0}{v(\mathbf{x})}\int d\boldsymbol{\phi}' = 0 \tag{5.5.12}$$

A diffusion approximation to the original random flight process can now be estimated. As defining properties of this approximation we choose

- The approximation must have local diffusivity $D = v^2\tau/2$ equal to the motility of the original process $\mathbf{x}(t)$.
- The approximation must share the steady-state Eq. 5.5.11 with the original process $\mathbf{x}(t)$.
- The stationary approximation must be time reversible as is the original stationary process $\mathbf{x}(t)$.

The diffusion process with these three properties is the one which has the Master equation

$$\frac{\partial p}{\partial t} = \nabla\cdot\left(\frac{\tau v}{n}\nabla(vp)\right) \tag{5.5.13}$$

To see that the stationary process with the Master equation (Eq. 5.5.13) is time reversible, notice that inserting $p = 1/v$ in equation Eq. 5.5.13 yields zero flux, not just a non-divergent flux. This, in the terminology of stochastic processes or statistical mechanics, is detailed balance (Gardiner, 2004) and implies time reversibility. This property makes the approximation unique.

5.5.2 Including locally exponential waiting times.

In this case the individual may be in one of two qualitative states, moving or pausing. A pause is followed by a move in a random direction; a run is followed by a pause. The duration of runs and pauses are locally exponential, *i.e.* a run or a pause is terminated at a rate which depends on current position, but is independent of the duration of the current run.

Let $\psi(\mathbf{x},\boldsymbol{\phi},t)d\mathbf{x}d\boldsymbol{\phi}$ be the number of individuals at time t which are moving in direction $\boldsymbol{\phi}$ near $\mathbf{x}$ and let $\rho(\mathbf{x},t)d\mathbf{x}$ be the number of individuals at time t which are pausing near $\mathbf{x}$. Then the Master equations governing these coupled processes are

$$\frac{\partial\psi}{\partial t} = -\nabla\cdot(v\boldsymbol{\phi}\psi) - \alpha\phi + \mu\rho \tag{5.5.14}$$

$$\frac{\partial\rho}{\partial t} = \alpha\int\psi d\boldsymbol{\phi}' - \mu\rho \tag{5.5.15}$$

where $\alpha = 1/\tau$ is the mean rate of transition from run to pause, and $\mu = 1/\sigma$ is the mean transition from pause to run. The steady-state solutions to these master equations are

$$\psi = \frac{1}{v} \qquad \text{and} \qquad \rho = \frac{\alpha}{\mu}\frac{1}{v} \tag{5.5.16}$$

which may be verified directly. The total concentration p_0 is then

$$p_0 = \int\psi d\boldsymbol{\phi}' + \rho = \frac{1}{v}\left(1 + \frac{\alpha}{\mu}\right) \tag{5.5.17}$$

The local diffusivity is (*cf.* Eq. 5.0.2)

$$D = \frac{v^2}{n}\frac{1/\alpha}{1+\alpha/\mu} \tag{5.5.18}$$

and the approximating advectiondiffusion equation is

$$\frac{\partial p}{\partial t} = \nabla \cdot \left(\frac{v}{n\alpha} \nabla \left(\frac{vp}{1+\alpha/\mu} \right) \right) \tag{5.5.19}$$

5.5.3 Random jumps in heterogeneous environments

Consider now the random jump model, in which the individual determines the parameters governing each the run at its onset. A consistent advection-diffusion approximation will be required to have the same initial growth of variance and the same mean dynamics. This specifies the approximation uniquely and leads to simple analysis since the expected displacement during a single jump is exactly 0. By induction it follows that the expected displacement during any number of jumps, or indeed during any fixed time interval is also 0. Thus in a heterogeneous environment the densities may develop skewness, but their mean remain constant. The advection-diffusion process with this property, and with the same initial growth of variance $2nD(\mathbf{x})t$ as the original random jump process, is governed by the Master equation

$$\frac{\partial p}{\partial t} = \nabla^2 \left(D(\mathbf{x}) p \right) \tag{5.5.20}$$

a well-known fact which can be verified by multiplying with the i^{th} component of vector $\mathbf{x}$ and integrating over $\mathbf{x}$. This result is consistent with the approximations for more general jump processes derived in (Okubo and Levin, 2001) using perturbation analysis of the Master equation, and in (Gardiner, 2004) using the backward equation; see also (Grünbaum, 1999).

The Master equation predicts a steady state density inversely proportional to $D(\mathbf{x})$. This, however, only holds in the diffusion limit where the characteristic time and length scales of the jumps approach 0 while keeping the local diffusivity $D(\mathbf{x})$ constant. It is a reasonable approximation when characteristic scales of the single jump are much smaller than those of environmental variation.

5.6 Individuals to populations

One of the major challenges facing marine ecologists is to link processes that govern the interaction of individual plankters with the environment and each other, to descriptions of ecosystem and population scale processes. In some sense this is the integrating link between laboratory studies of plankton behaviour (*e.g.* swimming, encounter, capture, escape, reproduction) and field observations of their distribution and function in the marine environment (*e.g.* blooms, swarms, growth, grazing, remineralization). For a large class of stochastic kinetic behaviours, parallels can be drawn between these biological processes and the physical formalism underlying the statistical mechanical description of thermodynamics.

This physical formalism has a long history stretching back to Maxwells kinetic theory of gases, and Einsteins treatment of Brownian motion in terms of the kinetic energy transfer between individual fluid molecules (*e.g.* Chapman, 1928). Direct transfer of these concepts to the problem of biological motility has been successfully implemented (*e.g.* Skellam, 1951; Okubo,

1980; Berg, 1992; Okubo, 1986; Schnitzer, 1993). However, the behaviour of biological agents is not as simple as that for molecules; the conservation of kinetic energy is not a constraint (organisms can stop and start using internal energy reserves) and the entropy of organism distributions can decrease (*i.e.* organisms can accumulate).

This difference between biological motility and the interaction of physical particles means that a direct transfer of methodologies must be done with caution. In particular, the parameterization of random motility in terms of "bio-diffusion" is incomplete. The advection-diffusion equation describing the continuum of organism distribution takes on different forms depending on the details of the motile behaviour.

In other cases, the analogy between physics and biology over-reaches itself. In applying the statistical mechanical treatment of Brownian motion in a potential field (Chandrasekhar, 1943), some models have introduce an "attractive force" acting on the organism (Okubo, 1972). Indeed the concept of an "attractive force" appears often in theoretical treatments of plankton motility. However, with the exception of light, there are very few physical parameters that marine pelagic plankton can perceive over a distance. Nearly all environmental information from the proximity of prey to the concentration of amino acids derives from a volume less than a body length in radius, much less than the average spacing between individuals and certainly less than the spatial dimension of a patch or a swarm (*cf.* chapter 6). Thus the concept of an "attractive force" is artificial, and while it can be used to mimic a macroscopic effect, it is unlikely to be a behavioural component of plankton motilities.

The qualitative analysis of motion has a long history in mathematical biology, and much has been written on motility, locomotion and migration where debate has often revolved around distinctions between taxis versus kinesis, continuous versus discrete random walks, biased versus unbiased random walks. All of these a valid topics of discussion, but seem to two fundamental and systematic issues.

- $\alpha-$ integration: That there is a whole spectrum of equally valid axioms in stochastic integration from which very nearly any macroscopic behaviour can be predicted.
- Diffusion is a composite parameter: How diffusion is partitioned between its temporal and spatial manifestations also impacts predicted macroscopic behaviour.

It appears that nothing can be proven with regards the micro-macro scale aspects of Brownian diffusion from first principals by appealing to mathematical formalism. Such aspects can only be pinned down by additional information.

6 Patchiness

The marine environment is far from homogeneous, and plankton is distributed un-evenly. Seen from space, the surface ocean appears as a mosaic of patches, filaments, tendrils and gyres. To a large degree, these features are made visible by plankton – either directly by colouring the water, or more subtly, by changing the reflectance properties of surface waters[1]. Closer examination at finer scales indicate that the un-even distribution is manifest wherever one looks. Fine-scale distribution patterns of dissolved material (nutrients, organic carbon), phytoplankton, bacteria and zooplankton all have a clumpy distribution. Patchiness is evident to a greater or lesser degree at all spatial scales and for all constituents of the marine ecosystem.

Some classic work in documenting and explaining fine-scale distribution of plankton was conducted over 50 years ago by Morrison Cassie (Cassie, 1963). I have a personal fondness for this work as it was conducted in Port Nicholson, (Wellington Harbour) on the shore of which I grew up. Cassie found that plankton is almost invariably distributed unevenly in space – a patchy distribution – in that plankton tend to cluster more than would be expected from a random (Poisson) distribution. This is termed "under-dispersed" as opposed to "over-dispersed" where distributions are more uniform than random. The precise difference between these distributions can be seen in various statistical measures of the distribution pattern such as the frequency distribution of inter-plankton separation distances, or the coefficient of variance. Over the last 50 years, research into patchiness in the marine environment has undergone sporadic bursts of activity covering 3 main topics; documenting its prevalence and structure, examining its underlying dynamics and identifying its impact on the ecology of marine organisms.

6.1 Patchiness and concentration

What does it mean to be patchy? Life in the ocean, after all, comes in packets. Organisms are discrete entities, so at some scale patchiness is inevitable. This may seem a bit semantic at first, but it speaks directly to the problem of measuring patchiness and its associated problem, the problem of measuring of concentration. These concepts are central to ecology in general – the study of the abundance and distribution of organisms – but both abundance and distribution presuppose some scale and some statistic of relevance.

For instance, consider a large volume V containing a corresponding large number N individual organisms. The concentration of organisms is thus $\rho = N/V$. But now sub-sample this original volume in n smaller equal volumes of $v = V/n$. The average number of organisms in each sub-volume is $\mu = N/n$, but what about the variance σ^2. For a uniformly random

[1] Oils produced by plankton accumulate in convergence zones, so suppressing capillary waves and sea-surface roughness

distribution (*i.e.* Poisson), there will be some expected variance due to random effects which turns out to be $\sigma^2 = \mu$. A useful measure of patchiness is thus the dispersion parameter:

$$\frac{\sigma^2}{\mu} \quad \begin{cases} < 1 & \text{under} - \text{dispersed, regular} \\ = 1 & \text{uniformly random, Poisson distributed} \\ > 1 & \text{over} - \text{dispersed, clumpy} \end{cases} \tag{6.1.1}$$

Patches appear over a whole spectrum of scales, from centimetres to kilometres.

The root-cause of this patchiness is almost invariably biotic in origin, due to growth and/or motility subsequently shaped by physical processes such as circulation patterns, fronts and turbulent stirring. Patchiness, in the case of motility, relates to body size, in that it reflects

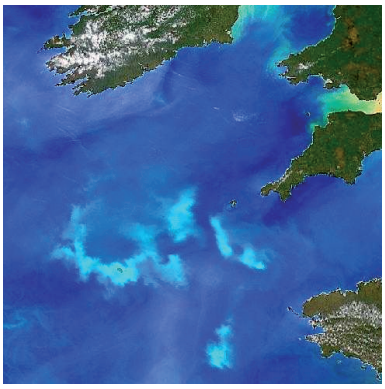


Figure 6.1: A bloom of coccolithophores in the Celtic Sea. Such blooms are common features around the world. They are formed by single celled protists, a few tens of μm's in diameter, that are armored with calcium carbonate plates. They are important in the carbon budget of the oceans, having been responsible in the past for laying down vast chalk deposits but are under threat by climate change and the lowering of the oceans' pH.

the environmental "grain" organisms experience. That is, the behaviour of organisms is most influenced by spatial variations corresponding in the scale to their ambit (*i.e.* the spatial scale over which they typically range). Variations smaller than this are seen as noise and averaged out, while larger scale variations are simply not perceived at all. Small organisms with correspondingly small ambits, concentrate on small scale features; larger organisms with larger ambits on larger scale features. Thus bacteria, flagellates and ciliates exhibit patchiness at the scale of centimetres reflecting the distribution of the solute patches and detrital aggregates they feed on. Copepods and krill on the other hand are distributed at the scale of 10's of metres or so, reflecting perhaps food abundance in subsurface layers, or as a counter measure to predation risk.

This of course cannot be the whole story. Phytoplankton blooms for instance, that are made up of micron scale organisms that hardly move at all, can be on the scale of 100's of kilometres (*cf.* Fig. 6.1). This of course reflects growth and the underlying distribution of available resources (nutrients or light). Look closely at a satellite image of a phytoplankton bloom and you will see finer and finer details in the distribution; the boundary is scalloped, coherent features such as filaments are evident in places, while other regions of the bloom look fuzzy and diffuse. Evidently, such patterns arise from chaotic stirring. As the bloom progresses, any number of physically driven flows such as Langmuir circulation, tidally rectified currents, frontal instabilities, inertial oscillations, and internal waves (all precursors and manifestations of oceanic turbulence) can interact to stretch, squeeze and deform the plume across a whole range of spatial scales.

Within the concept of grain, there is an important distinction to be made as to the mechanisms by which these spatial structures are created and maintained, namely aggregative processes and re-distribution processes. Aggregation implies an increase in concentration; a

generation of variance if you will, and necessarily requires a biological factor at its source (*e.g.* growth, migration, swimming in concert with convergent flows, or self-organized aggregates such as swarms and schools). Re-distribution processes on the other hand are generally physically driven; chaotic stirring generates small scale grain from larger scales –the cascade of variance (Mackas et al., 1985). That is, chaotic stirring dissipates patches by breaking them into smaller and smaller bits. The realized grain of the marine environment is due to a continual generation, cascade and dissipation of variance driven by a variety of biological and physical processes operating across a broad range of spatial and temporal scales.

6.2 Stirring, birth, death and inertia

To illustrate how turbulent stirring and biological processes interact to generate patchiness, I will introduce 2 elegant models that have helped shape my understanding of patchiness.

The first contains only 3 elements; organisms give birth and die in a stirred fluid (Young et al., 2001). A clustered distribution emerges from birth-death processes by the simple fact that organisms are necessarily born close to their mother (and hence siblings and cousins *etc.*), but in general die randomly in space. This is clear enough, but the effect is almost impossible to capture in advection-diffusion-reaction equations, the most readily accessible and most often used means of modelling marine ecosystems.

Suppose for instance there is a population of organisms for which the birth rate is α, and death rate is μ. The population is subject to some mixing defined by the spatially uniform diffusivity D. In a continuum view, the concentration of organisms $C(\mathbf{x}, t)$ thus is given by

$$\frac{\partial C}{\partial t} = D\nabla^2 C + (\alpha - \mu)C \tag{6.2.1}$$

subject to some initial condition and boundary conditions (which I assume here to be a zero flux condition). This is most readily seen in terms of a 2 D population, but can be extended to 3 D.

For the case where births and deaths occur at the same rate, $\alpha = \mu$, any inhomogeneity in the population distribution is eventually smoothed out. Put another way, if the distribution is uniform initially (*i.e.* $C(\mathbf{x}, t = 0) = C_0$), it will always be uniform (*i.e.* $C(\mathbf{x}, t) = C_0 \quad \forall t$).

The same case can be examined using discrete particles that reproduce, die and diffuse in a random way. A simple numerical experiment (Bill Young's so-called Brownian Bug model) shows the clustering effects over successive generations (Fig. 6.2).

In sharp contrast to the continuum description, an individual-based rendering of the same basic population dynamics generates a distribution pattern that is far from uniform. The birth-death component of the problem is a Galton-Watson process – a statistical process describing the probability that a lineage will die out after a number of generations given a particular reproduction rate.

The name derives from Francis Galton (1822 – 1911) a cousin of Darwin who, amongst a great many other scientific pursuits[2] became interested in the rate at which family names in Britain disappeared. It was the Reverend Henry William Watson (1827 – 1903) who came up with a solution and together, they wrote "On the probability of extinction of families", published in

[2]He wrote many articles and several books on fingerprinting, heredity, meteorology, statistics and introduced the concept of the daily weather map in newspapers. He also held many odd views, promoting eugenics for instance where his name lingers in the Galton Institute.

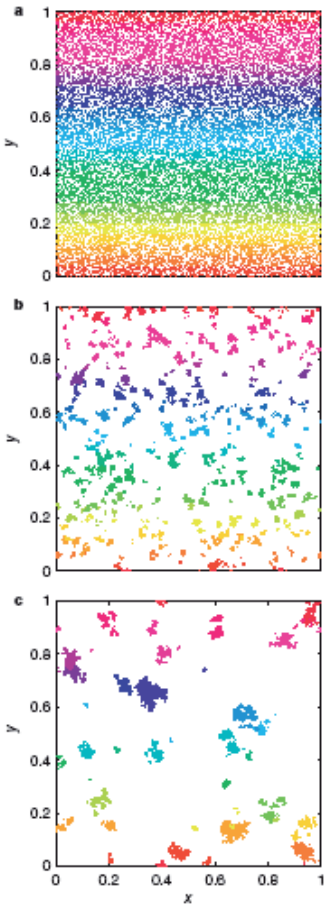


Figure 6.2: A numerical simulation of the evolution of clusters by a birth-death-diffusion process. Organisms are initially distributed randomly in space (a Poisson distribution). The birth rate and death rate are equivalent. N_d of a total N organisms are randomly selected for demise after which an equal number of the surviving $N - N_d$ are randomly selected to reproduce. Daughter organisms are placed at the same location as their mother, and given the same color code. All organisms then undergo diffusion via a discrete random walk: a uniformly distributed random angle $\theta_i \in [0, 2\pi]$ is chosen for each organism, and their position is updated according to $\mathbf{x}_i \leftarrow \mathbf{x}_i + v\delta\mathbf{r}_i$ where $\mathbf{r} = \cos\theta_i\hat{\mathbf{x}} + \sin\theta_i\hat{\mathbf{y}}$, v is a diffusive velocity scale, and δ is the time step. Diffusivity is thus given by $D = v^2\delta/4$, appropriate for 2 dimensions and for uniform Brownian run lengths.

Figure redrawn from Young et al. (2001)

1874. In the simplest case, suppose that at each generation, an organisms (or father with a distinguished family name) has ξ offspring drawn from a Poisson distribution with intensity λ. The probable number of survivors from a single antecedent after n generations is given by the iterative relationship $x_{n+1} = e^{\lambda(x_n - 1)}$. Galton-Watson processes are useful descriptors not just for geneology, but also for evolutionary ecology (*e.g.* neutral theory), and branching processes (*e.g.* percolation theory, epidemiology, nuclear chain reactions).

Invariably, birth-death (and likely other biological processes) lead to pair-correlations in the distribution of organisms. That is, they are not Poisson distributed with the result that continuum descriptions of the distribution of organisms is not well posed. This has potentially far reaching consequences for the fidelity of traditional advection-diffusion-reaction equations in simulating marine ecosystems.

The birth-death process proposed above indicates an important patch generating mechanism, but it is played out in a wider context. The basic premise is that conditions for growth, reproduction and mortality are isotropic everywhere. But these parameters in themselves are patchy. Nutrients are up-welled on large scales and remineralized at small scales with the effect that the potential for growth is patchy. Realized growth requires that such patches are colonized – which may be a haphazard process. Reproduction depends on mature organisms meeting in the rarified reaches of the pelagic ocean. Finally, predators do not pick off random individuals uniformly in space, but sweep through communities in a spatially connected

way – eating holes in its prey distribution and giving it a swiss-cheese mosaic. Abundance and distribution of organisms depend in part on the trophic interactions within the marine ecosystem.

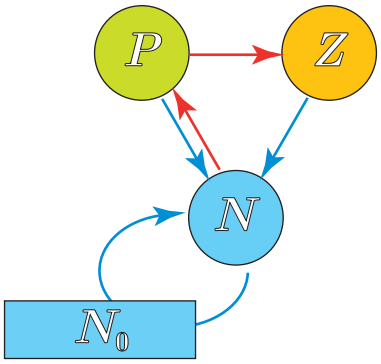


Figure 6.3: A hybrid between individual-based and continuum-based models is to consider fluid elements that carry with them a drifting micro-ecosystem. In a simple rendering, fluid elements are advected through space (2 dimension is easiest to conceptualize) is some quasi-random manner. Within each fluid element, a NPZ (*i.e.* nutrients – phytoplankton – zooplankton) model tracks the dynamics of the food web. Phytoplankton grow at a rate limited by nutrient availability, and are grazed by zooplankton. Zooplankton grow at a rate determined how much they graze. Nutrients are recycled every time an organism dies, and is slowly exchanged with a back ground pool.

In order to gain insight into these processes Ed Abraham (Abraham, 1998) developed a relatively simple model (*cf.* Fig. 6.3), incorporating NPZ dynamics into a stirred ocean model.

As in all such modelling exercises, elegance and simplicity far outweigh over-detailed realism and complexity. The dynamics of the NPZ model can be written as:

$$\frac{dN}{dt} = \gamma(N_0 - N) \tag{6.2.2}$$

$$\frac{dP}{dt} = \alpha P\left(1 - \frac{P}{N}\right) \tag{6.2.3}$$

$$\frac{dZ}{dt} = P(t-\tau)Z(t-\tau) - \mu Z^2 \tag{6.2.4}$$

where γ is the nutrient relaxation rate, α is the growth rate of phytoplankton limited by a carry capacity, τ is a zooplankton generation time scale, and μ represents zooplankton mortality. Equations similar to these (albeit with more compartments, a great many more parameters, sub-models and processes) are at the heart of nearly all marine ecosystem models. Interpreting these in a Lagrangian sense, we follow the dynamics of N, P and Z as fluid elements travel about in space. Specifically, if $\mathbf{x}^\bullet(\mathbf{x}_0, t)$ is the unique trajectory of a fluid element specified by its location $\mathbf{x}^\bullet = \mathbf{x}_0$ at $t = 0$, then N implies $N(\mathbf{x}^\bullet(\mathbf{x}_0, t))$. Likewise for P and Z. The variable N_0 however is an Eulerian field, fixed in space and not advected by the flow. This may for instance be interpreted as a deep-water pool of nutrients that slowly up-wells into surface waters.

To complete the model, some simulation of turbulent stirring is required. Abraham (1998) used a seeded-eddy model; horizontal eddies of different size are set to spin so that the kinetic energy spectrum conforms to a given wave-number spectrum – $k^{-5/3}$ for instance – but it could be any empirically or theoretically derived spectrum. The eddies advect not just passive tracers but each other. While the basic seeding process is fairly straightforward, I've never figured out how to efficiently implement boundary conditions. An alternative is to use a renovating wave model (Young et al., 2001) introduced in Sec. 4.3.2.

By what ever means the physical stirring is implemented, it becomes clear that biotic processes that have a long time scale have a more fine scale variance than those with faster time

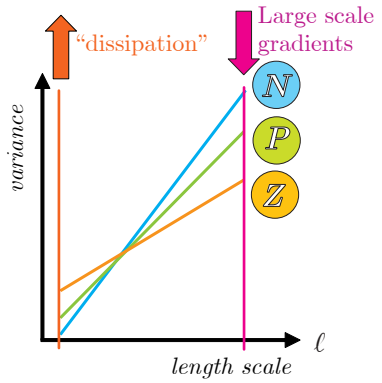


Figure 6.4: The variance spectrum of distribituions become whiter for higher trophic levels.

scales. The nutrients (carrying capacity) very nearly follow the distribution predicted for an ideal tracer via the cascade of variance. Higher trophic levels integrate abundance of their resources within the moving water and can thus be thought of as having a greater biotic inertia. What does this mean for evolving distribution patterns? In the same way that the mean separation distance between two particles increases with time, one can think of running the clock backwards so that two particles close together now, were separated by increasingly greater distances in the past. The variance of some property contained in near neighbour fluid elements is greater for properties with long integral time scales (Fig. 6.4). That is, higher tropic levels display higher small scale variance (*i.e.* greater small scale patchiness) than lower trophic levels which follow closer to an inert tracer spectrum. This pattern is well documented in nature (Mackas et al., 1985).

6.3 Encounter-mediated kinesis and self seeding patches

Pair-correlations, and thus patchiness can be brought about by means other than localized growth. Many planktonic organisms are motile to a greater or lesser degree. The overarching rational for motility is connected with the primary activities of all organisms: the search for prey and mates and the avoidance of predators. Depending on the behavioural response an encounter elicits, the patchy distribution of targets acts as an attractant or repellant – imprinting spatial patterns on the distribution of searchers. When some motile behavior is elicited by encounters with con-specifics, spatial patterning can emerge without any external cue.

Here I am interested in the formation of so-called micro-scale patches (cm to 10 m scale) where individual behaviour is believed to be an important biological driver (Folt and Burns, 1999). Indeed, many studies have shown that zooplankton change their swimming behaviour in encountering prey (*e.g.* Buskey, 1984; Tiselius, 1992; Menden-Deuer and Grünbaum, 2006) and mates (*e.g.* Buskey, 1998; Bagøien and Kiørboe, 2005a,b). The rationale for focussing on motility and patchiness stems from the observation that in pelagic systems there are very few physical cues other than the patchiness of the environment itself to serve as attractors for aggregation. Furthermore, growth rates of zooplankton are slow compared to the time scale of mixing, so that patches are likely to represent animals which were born and grew up at some different location.

Encounter-mediated kinesis in an otherwise random walk motility of zooplankton can induce aggregation. This behaviour can be directly related to the sensory and swimming abilities of zooplankton. When the encounter-mediated kinesis (*e.g.* encounter-pause behaviour) is

elicited by con-specifics, aggregation can occur in the absence of any extra-specific cue in a process of self-seeding patchiness. In the natural environment, clustering is counter-acted by turbulent dispersion of patches. This is a scale dependent process which has a structuring effect on the patchiness of pelagic organisms as a function of

- the biological process supporting patch formation (growth, migration or quasi-random motility)
- the size-dependent sensory and swimming abilities of the organisms involved and
- the nature and function of the "attractor" (fixed in space or drifting, self-sustained or external).

When balanced against scale dependent turbulent dispersion, different biological processes lead to different preferential patch sizes.

One unsatisfactory aspect of works to date examining aggregation by means of motility, is that they fail to identify specific sensory mechanisms that allow zooplankton to form and remain in patches. This is of fundamental concern when one considers the nearest neighbour distances represented by mean concentrations (100's of body lengths), compared to the limited sensory ability of most zooplankton (a few body lengths at most). Most previous works have invoked an "attractor" towards which zooplankton swim. However, with the exception of light, there are very few environmental signals in pelagic systems that zooplankton can detect over long distances, and which may serve as an attractor. Patches of prey and chemicals have been hypothesized as attractors, but this simply postpones the question as to what processes serve to form and maintain these lower tier patches.

Here I will explore the concept that the exchange of information can be the "glue" that holds patches together.

Consider for instance a hydromechanically sensitive (rheotactic) searching organism (denoted i), detecting a target organisms (denoted j). The encounter rate Z_{ij} can be written

$$Z_{ij} = \pi R_{ij} C_j \left(v_i^2 + v_j^2 + 2w_{ij}^2\right)^{1/2} \tag{6.3.1}$$

where the target organism j has a local concentration C_j, R_{ij} is the distance at which i can detect j, and v_i and v_j represent the mean speed of the randomly directed swimming of each species. The effect of turbulence is incorporated through w, the turbulent velocity scale. Scale analysis of the hydromechanical signal generated by self-propelled organisms, and the sensory ability of copepods suggests that the reaction distance is given by

$$R_{ij} = a_j \left(\frac{v_j}{s_i}\right)^{1/2} \tag{6.3.2}$$

(Visser 2001), where a_j is the size of the target organism (j) swimming with speed v_j, and s_i is the threshold velocity of the hydromechanical disturbance that searcher organism (i) will react to. Notably, the reaction distance is a function of both the searcher's detection ability (s_i) and target's conspicuousness (a_j, v_j).

Within this context, I will examine information exchange and simple motile behaviour that can lead to patch formation, and whether this behaviour is sufficient to maintain patches within a turbulent environment.

6.3.1 Random motility and "bio-diffusion"

As discussed extensively in chapter 5, the quasi-random motility of an organism may be treated as a diffusive-like process, that can be locally parameterized by a diffusivity $D(\mathbf{x}, t)$. In general, the parameters controlling this bio-diffusion (*e.g.* speed, mean free path, pause interval) can vary from place to place ($\mathbf{x}$) and time to time (t) giving rise to potential aggregating effects. In what follows, I will assume that the details of the motility are such that the standard Itô stochastic integral can be used (*cf.* Sec. 5.4). Further, the organisms may be under the influence of physically mediated diffusion; turbulence for instance, or Brownian motion proper. Writing the physical diffusivity as K, the Fokker-Planck equation describing the describing the concentration C of similar organisms, is given by

$$\frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{J} = \nabla \cdot \{K \nabla C + \nabla (DC)\} \tag{6.3.3}$$

where K is the physical diffusivity and $\mathbf{J}$ is the flux.

Note that if D is spatially uniform, both physical and bio-diffusion lead to a dispersion of organisms. However, if D is not spatially uniform, random motility can lead to an accumulation of organisms in regions where D is low. The flux can be written as

$$\mathbf{J} = -(K + D)\nabla C + C \nabla D \tag{6.3.4}$$

and is composed of 2 parts. Firstly, there is a diffusive contribution with flux going from high to low concentration at a rate determined by $(K+D)$, and secondly, there is an advection-like contribution going from regions of high to low bio-diffusion at an apparent advection velocity ∇D.

6.3.2 Social aggregation by motility

To illustrate this advective-diffusive nature, consider an important case of bio-diffusion where the motility D is a function of the organisms own concentration C. That is, the random walk parameters are functions of the proximity of con-specifics. To explore this case, assume a general power relationship of the form $D = aC^n$ where a is a constant specific for the power n. Substituting into Eq. 6.3.4, this gives

$$\mathbf{J} = -(K + a(1 + n)C^n)\nabla C = Q \nabla C \tag{6.3.5}$$

The evolution of small fluctuations in the concentration field can be determined from the characteristics of Q. Specifically, when $Q < 0$ everywhere in space, the flux will always be directed down gradients of C. Thus any small perturbation in the concentration distribution will always decay. Because K, a and C are all positive quantities, Q will always be negative when $n > -1$. When $n = -1$, there is no net effect from motility; the diffusive contribution is exactly balanced by the advective contribution and again small perturbations in C will decay. However, if Q changes sign at some locations, (*i.e.* $Q = 0$ somewhere) then some perturbations will grow. That is, social aggregation can take place. A necessary condition for this to be satisfied is $n < -1$.

Fig. 6.5 illustrates how this can lead to the growth of perturbations, depending on the coefficient n, and a critical population density C^*. For the case where $C^* < C_{min}$ (Fig. 6.5.c), physical diffusion dominates everywhere, and all perturbations eventually decay. When $C^* > C_{max}$ (Fig. 6.5.b) motility dominates everywhere and peaks in the concentration act as local attractors and intensify. For $C_{max} > C^* > C_{min}$ (Fig. 6.5.a), motility dominates for low concentrations, and physical for high concentrations. Peaks are eroded while troughs deepen, and there are local attractors at $C = C^*$.

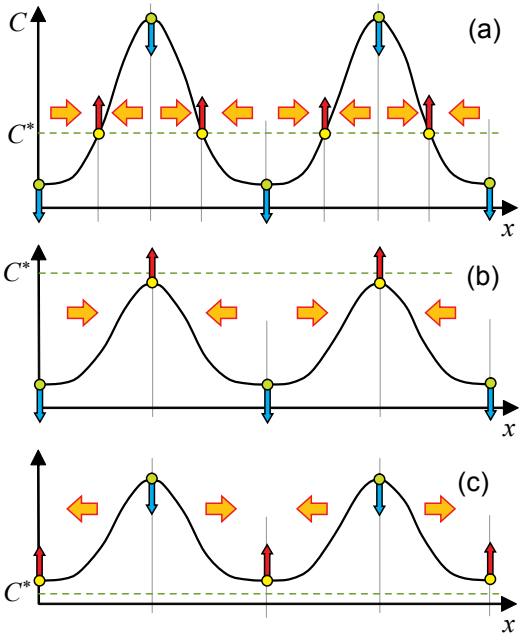


Figure 6.5: The aggregation/dispersal of organisms under the influence of density dependent motility $D = aC^n$, where C is the concentration of organisms, and K is the spatially uniform physical diffusivity. Auto-aggregation is possible when the power coefficient $n < -1$. For instance, for $n = -2$, $Q = aC^{-2} - K$, which equals zero when $C^* = (a/K)^{1/2}$. (a) Points where $C = C^*$ act as local attractors, peaks are eroded and troughs deepen. (b) When $C < C^*$ everywhere, peaks intensify and auto-aggregation occurs. (c) For $C > C^*$ everywhere dispersion dominates and the distribution becomes smooth.

6.3.3 Encounter-pause motility

Let us consider now the encounter process in more detail. Consider an organism with a reaction radius R_{ii} swimming at a constant, randomly directed speed v_i in a field of con-specifics of concentration C_i. Starting from a given location, the average time it must swim before encountering a con-specific is

$$\tau_s = \frac{1}{Z_{ii}} = \frac{1}{\beta_{ii} C_i} = \frac{1}{\pi R_{ii}^2 (2v_i^2 + 2w^2) C_i} \tag{6.3.6}$$

where Z_{ii} is the con-specific encounter rate and w is the turbulent velocity scale given at distance R_{ii}. On encountering a con-specific, the organism pauses for a time interval τ_p before turning in a random direction and resuming its swimming at speed v_i. The total mean time between encounters is thus

$$\tau = \tau_s + \tau_p$$

and the organism's mean speed between encounterers is

$$\bar{v} = \frac{\tau_s}{\tau_s + \tau_p} v_i \tag{6.3.7}$$

Thus the organism's effective motility is given by

$$D = \frac{\bar{v}^2 \tau}{3} = \frac{v_i^2}{3} \frac{\tau_s^2}{\tau_s + \tau_p} = \frac{v_i^2}{3\beta_{ii} C_i (1 + \beta_{ii} C_i \tau_p)} \tag{6.3.8}$$

Examining the asymptotic values of this motility, if the organism does not pause ($\tau_p = 0$), then $D = v_i^2/(3\beta_{ii} C_i)$, and motility has no effect on the redistribution of organisms. However, when the pause time becomes large compared to the encounter time $\tau_p \gg \tau_s$, then $D \approx v_i^2/(3\tau_p \beta_{ii}^2 C_i^2)$, which can enhance patches of organisms.

Some general properties of encounter-pause motility can be gleaned from its associated flux within a diffusive fluid. Substituting Eq. 6.3.8 into Eq. 6.3.4 yields

$$\mathbf{J} = Q\nabla C = \left(-K + \frac{v_i^2 \tau_p}{3(1+\beta_{ii} C_i \tau_p)^2}\right) \nabla C \tag{6.3.9}$$

Maximum motility occurs when encounter rates are low. As C_i becomes small, the condition for aggregation becomes $v_i^2 \tau_p / 3 > K$. That is, stronger swimming and a longer pause time lead to a greater ability to aggregate under similar diffusive backgrounds. In a manner similar to the previous section, a critical concentration C^* can be defined where $Q = 0$. This is given by

$$C^* = \frac{v_i (\tau_p / 3K)^{1/2} - 1}{\beta_{ii} \tau_p} \tag{6.3.10}$$

For $C > C^*$, perturbations in the organism concentration field decay, while for $C < C^*$, perturbations are grow and aggregation can occur. Scaling of C^* follows $R_{ii}^{-2} \tau_p^{-1/2}$ and appears to be largely independent of swimming ability. Thus organisms with large reaction distances have low critical concentrations.

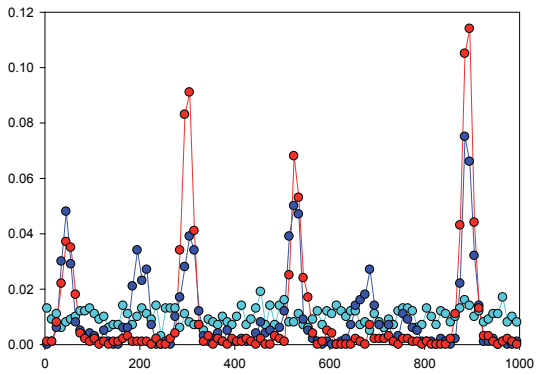


Figure 6.6: The aggregation of an encounter-pause organism starting from an initially uniformly random distribution (cyan dots). Parameters: mean concentration 0.01 ind/l, swimming speed v_i = 1 cm/s, detection distance R_{ii} = 1 cm, pause interval $\tau_p = 10$ s. Physical diffusivity is set to $K = 10^{-1}\text{cm}^2\ \text{s}^{-1}$. Given these parameters, the critical concentration below which auto-aggregation can proceed is $C^* = 1.3\ \text{L}^{-1}$. Distribution after 1.2 hours (blue dot) and 2.4 hours (red dot) show an increasing degree of aggregation.

This aggregation process is illustrated in Fig. 6.6. Parameters are chosen as representative for an adult copepod. Initially, 1000 particles are randomly distributed over 10 m. After 1 hour, small scale patches have appeared with length scales of about 60 cm and peak concentrations up to 5 times the mean concentration. After 2.4 hours, the patches are more discrete, with length scales about 70 cm, and peak concentrations up to 9 x the mean concentration. These numerical experiments show that encounter-pause kinetic behaviour can induce significant patchiness in the distribution of planktonic organisms. The scale and intensity of patches depends on the relative intensity of the kinetic behaviour compared to the eroding effects of physical diffusive processes. From the scale of individual plankters and up, these processes are governed by turbulence.

6.3.4 Scale dependent dispersion

A smooth transition of scale dependent turbulent diffusivity across the inertial subrange to eddy diffusivity is proposed here following the formulation:

$$K(\ell) = \alpha \left(\frac{L^4 \ell^4}{L^4 + \ell^4}\right)^{1/3} \tag{6.3.11}$$

where ℓ is the scale of a patch and L is the integral length scale. There is no specific reason for choosing this form other than it is convenient, has the appropriate asymptotic limits, and is smoothly continuous across all scales.

Examining motility of the form of Eq. 6.3.8 in the context of scale dependent turbulent dispersion, following Eq. 6.3.4 the flux of organisms can be written as

$$\mathbf{J} = Q\nabla C = \left(-K(\ell) + \frac{v_i^2 \tau_p}{3(1+\beta_{ii} C_i \tau_p)^2} \right) \nabla C \tag{6.3.12}$$

For a vanishingly small patch(*i.e.* , $\ell \to 0$), turbulent dispersion becomes negligible (molecular) so that motility will always tend to dominate and lead to aggregation. At any given length scale, the critical concentration is given by

$$C^*(\ell) = \frac{1}{\beta_{ii}\tau_p} \left(\left(\frac{v_i^2 \tau_p}{3K(\ell)} \right)^{1/2} - 1 \right) \tag{6.3.13}$$

That is, when the local concentration is less than C^*, motility dominates and aggregation occurs (Fig. 6.7).

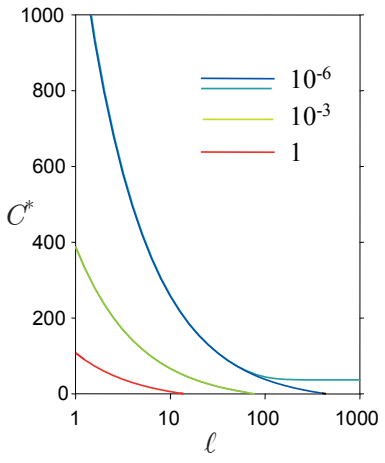


Figure 6.7: The critical concentration for a motile copepod (v_i = 1 cm/s, R_{ii} = 1 cm, τ_p = 10 s) under spatially dependent turbulent dispersion as predicted by Eq. 6.3.13. Curves are plotted for 3 different turbulent dissipation rates (units $\mathrm{cm}^2/\mathrm{s}^3$) with integral length scale $L = 10^5$ cm, and an effective Richardson's constant of 10^{-1}. At a low turbulence dissipation rate, $\varepsilon = 10^{-6}$ $\mathrm{cm}^2/\mathrm{s}^3$, patches of up to 70 individuals per litre can be maintained by auto-aggregation at scales up to 50 cm. Even at moderately high levels $\varepsilon = 10^{-3}$ $\mathrm{cm}^2/\mathrm{s}^3$, patches of up to 8 individuals per litre can be maintained at the same scale. Only when dissipation rates become very high, $\varepsilon = 1$ $\mathrm{cm}^2/\mathrm{s}^3$, do patches at scales 50 cm or greater become totally dominated by dispersion.

Examining Eq. 6.3.13 the maximum scale for which auto-aggregation can occur is given by

$$\ell^* = L \left(\left(\frac{3\alpha}{v_i^2 \tau_p} \right)^3 \varepsilon L^4 - 1 \right)^{-1/4} \tag{6.3.14}$$

That is, when dissipation rate increases, the critical length scale decreases as $\varepsilon^{-1/4}$. Conversely when the swimming speed or pause interval increase, the critical length scale increases as $(v_i^2 \tau_p)^{3/4}$.

6.3.5 Growth, motility, migration and turbulent dispersion

Fifty years ago, scaling arguments and analyses first appeared that gave estimates of the size of plankton patches as a balance between plankton growth and dispersion (Skellam, 1951; Kierstead and Slobodkin, 1953). In deference to the authors of these pioneering studies, the scale at which a balance between growth and diffusion is struck, is termed the KISS scale.

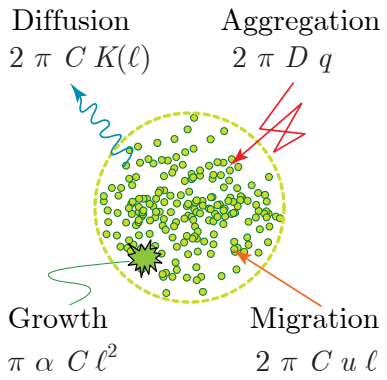


Figure 6.8: The critical size of a patch can be determined through a balance between effects that promote accumulation (growth, aggregation and migration) and dispersive effects (notably diffusion, but it may also include grazing). Accumulation and dispersion are in general dependent on the patch size ℓ.

It is of some interest to compare this scale with similar scales driven by other aggregating processes such as migration and motility, not only in the absolute scale dimensions involved, but also the functional dependence of these scales to both biotic and abiotic parameters.

Growth vrs dispersion: Consider a 2 dimensional patch or radial dimension ℓ. The net rate of increase of plankton within this patch depends on the specific growth rate α, the concentration of plankton C and the area of the patch $a = \pi\ell^2$. The rate of loss of plankton due to dispersion depends on the diffusive velocity $w = K/\ell$ (K being a uniform diffusivity) across the boundary of the patch (Fig. 6.9). The length of the boundary is $2\pi\ell$, so that the net outward flux is $2\pi CK$. When ratio of these, $G = \frac{1}{2}\alpha\ell^2/K$ equals 1, diffusion and growth balance, and the scale at which this happens $\ell^* = \sqrt{(2K/\alpha)}$ is termed the KISS scale (Okubo, 1980).

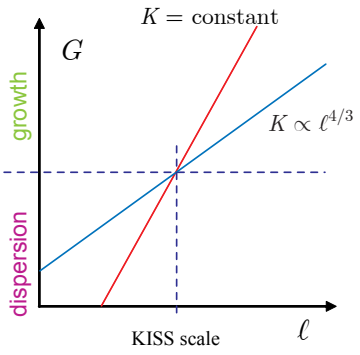


Figure 6.9: Growth dominates ($G > 1$) for scales larger than ℓ^* for both uniform and turbulent dispersion.

Patches at scales smaller than ℓ^* will decrease in concentration and eventually disappear. Larger patches however, will continue to grow in concentration until such time as the resources for growth (e.g. light or nutrients) become limiting, cell begin to sediment out due to increased coagulation, or the highly concentrated patch attracts the attentions of grazers.

Incorporating the scale dependence of turbulent diffusion changes the functional form of G, ($G \propto \ell^{2/3}$ for $\ell < L$, $G \propto \ell^2$ for $\ell > L$) but it remains an increasing function of ℓ in both cases (Fig. 6.9). Thus, the KISS scale represents the smallest scale at which a patch of reproducing plankton can exist in a turbulent environment also.

Migration vrs dispersion: In a direct analogy, a comparison of migration and dispersion (*e.g.* Okubo 1980) is given in Fig. 6.10. Here, migration signifies directed swimming correlated to an external cue. In this case, migration could be due to an encounter-pause motility cued to

a patchy prey field (*e.g.* growing plankton governed by KISS dynamics). Specifically, if the prey has peak concentrations of the order C_j and patch dimension ℓ_j, then the motility is given by

$$D = \frac{v_i^2}{3\beta_{ij}C_j(1+\beta_{ij}C_j\tau_p)} \tag{6.3.15}$$

and the migration velocity is given by

$$\mathbf{u} = \nabla D = -\frac{D}{\ell_j}\frac{\tau_s + 2\tau_p}{\tau_s + \tau_p}\hat{\mathbf{r}} \tag{6.3.16}$$

τ_p in this case can easily be interpreted as handling time. The negative sign indicates that it is directed towards high concentration C_j.

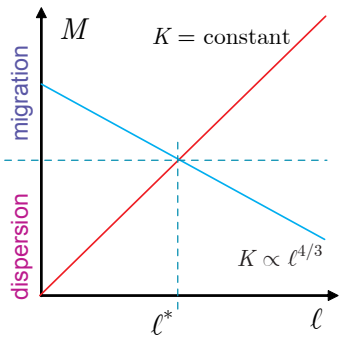


Figure 6.10: Migration dominates ($M > 1$) for scales larger ℓ^* for uniform dispersion, but smaller than ℓ^* for turbulent dispersion.

Unlike growth, both of the processes are governed by the flux of organisms across the patch boundary. The ratio of the inward migration flux ($2\pi\ell|\mathbf{u}|C_j$) to the outward diffusion flux ($2\pi CK$) is given by $M(\ell) = u\ell/K(\ell)$. A balance between migration and dispersion leads to different meanings of the balance point ℓ^* depending on the form of diffusivity. For spatially uniform diffusivity (*e.g.* turbulence at scales larger than L), ℓ^* represents a *minimum* patch size. However, for scale dependent diffusivity (*e.g.* turbulence at scales less than L), ℓ^* represents a *maximum* patch size.

Social aggregation vrs dispersion: To complete the sequence, consider the balance between aggregation and dispersion where in this context aggregation signifies the effect of encounter-pause motility cued to con-specifics. In this case, the organisms' motility is given by Eq. 6.3.8, and the net flux is given in Eq. 6.3.12. The ratio of inward to outward fluxes is thus given by

$$A(\ell) = \frac{J_{in}}{J_{out}} = \frac{D}{K(\ell)}\frac{\tau_p}{\tau_s + \tau_p} \tag{6.3.17}$$

In this case, the balance point ℓ^* where dispersion and motility are equal and opposite ($A = 1$) represents a maximum patch scale for scale dependent turbulent dispersion (Fig. 6.11). For uniform dispersion, there is no preferential scale; either all scales are possible or none are.

Scaling and turbulence: Thus, growth, migration and auto-aggregation each have different structuring effects on plankton patchiness. This is illustrated in Fig. 6.12 and shows how the 3 different processes are effected by shifting turbulent dissipation regimes as formulated in

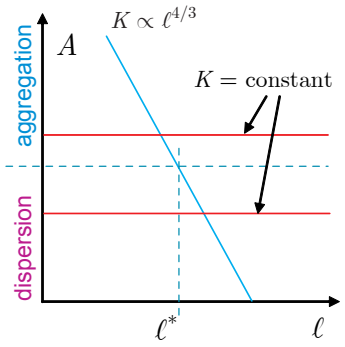


Figure 6.11: Auto-aggregation dominates ($A > 1$) for scales smaller than ℓ^* for both uniform and turbulent dispersion.

Eq. 6.3.11. The integral length scale L is set to 1 km, commensurate with turbulent eddies at the tidal excursion length scale. For the turbulent diffusivity found at this scale, and for the dispersion rates listed, an appropriate Richardson's constant is $\alpha' = 10^{-2}$. $G(\ell)$ is an increasing function of ℓ over all scales. Thus ℓ^* where $G(\ell^*) = 1$ represents a minimum allowed scale, and increasing turbulence and/or decreasing growth rate increase the value of ℓ^*. In comparison, $M(\ell)$ has a minimum near $\ell = L$. For strong swimmers and at weak turbulence, $M > 1$ at all length scales so that any patch size can exist. However, as turbulence

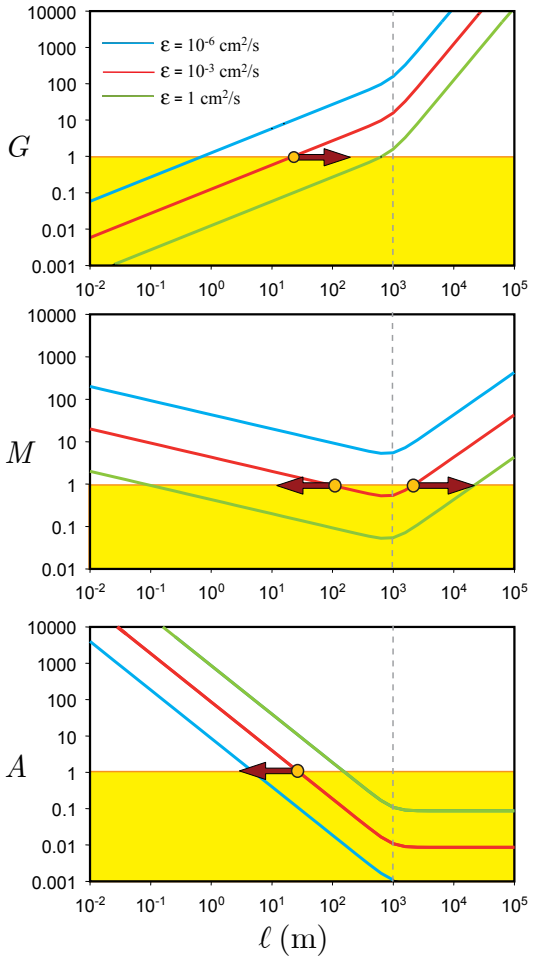


Figure 6.12: The ratio of aggregative processes to turbulent dispersion for growth (G), migration (M) and auto-aggregation (A). Where these ratios are greater than 1, aggregation dominates and patches can be maintained. Turbulent dispersion is scale dependent following $K(\ell) \approx \alpha' \varepsilon^{1/3} \ell^{4/3}$ for $\ell < L$, and is constant for $\ell > L$. The integral scale $L = 1$ km, and for which the appropriate Richardson's constant is $\alpha' = 10^{-2}$. The accumulation–dispersion ratios are calculated for 3 levels of turbulent dissipation rate representative of shelf seas; $\varepsilon = (1, 10^{-3}, 10^{-6})$ cm^2/s^3.

increases or swimming ability decreases, a dispersion-dominated range bracketing L appears and grows as turbulence increases even further. The critical scale ℓ^* where $M(\ell^*) = 1$ has two values representing both a minimum and maximum scale. The scale of patches that can be maintained by migration diverges into scales both larger and smaller than L. Finally $A(\ell)$ is a decreasing function of scale where and the critical scale ℓ^* where $A(\ell^*) = 1$ represents a maximum allowed scale. Thus, under the same turbulent conditions, strong swimmers such as copepods can form large patches while at the same time weak swimmers such as bacteria can form small patches.

The formation and maintenance of dense plankton accumulations can be mediated by two general processes: growth and active aggregation through swimming. These aggregating processes play out within a turbulent environment where large scale distributions are continually broken down into finer grained distributions through the turbulent energy cascade. Within this context, growth appears to be an important driver of planktonic patchiness (Abraham 1998). However, it is unlikely that growth is the only driver for observed zooplankton patches. While anecdotal reports from nature exist, active aggregation has certainly been observed under laboratory conditions. Furthermore, active aggregation can be inferred from field observations which show that patches are often mono-specific indicating they arise through species specific and/or social behaviour. Finally, even with limited swimming ability, zooplankton are far from passive tracers and can travel considerable distances over their lifetime.

Active aggregation can be mediated by circulation patterns. In particular, vertical swimming in concert with horizontal convergence is effective in creating elevated concentrations of plankton associated with internal waves, Langmuir cells and fronts. Other forms of active aggregation depend on cued behaviour attuned to a physical attractor. Such an attractor could be drifting debris, a shaft of light shining through mangrove roots, a particular shade of sand on the seafloor, a coral head or a specific light level below the surface. While some of these attractors certainly exist, a large body of work proceeds on the assumption that aggregation is attractor driven. The case for attractors breaks down particularly in the pelagic where there are very few physical cues. This work seeks to circumvent the need for a supposed attractor by examining the dynamics induced by density dependent motility arising through the encounter rate of rheotactic swimming organisms. Density dependent motility is modelled here as an encounter-pause process. That is, an organism swims until it encounters a con-specific whereupon it pauses for a given period before setting off in a new random direction. It appears that motility of this form is effective in generating mono-specific patches of zooplankton many times background concentrations, up to metres in scale and within time periods of a few hours.

When balanced against scale dependent turbulent dispersion, different biological and behavioural processes driving aggregation lead to a structuring of plankton distribution patterns. Each of the three processes considered here: growth, migration attuned to an external attractor, and auto-aggregation due to density-dependent motility, each have distinct scale dependence. While cast in terms of horizontal 2D patches, the same scaling laws would be found in a 3D analysis. Growth depends on the patch volume, migration on its surface area, and diffusion, both through physical diffusion and motility, depends on the linear patch dimension. Thus the derived scaling laws are relatively general: growth leads to large patches, auto-aggregation to small patched and migration to both small and large patches separated by a range encompassing the integral length scale.

6.4 Lévy walks

Over the last couple of years, Lévy walks have become a rather fashionable topic in marine ecology as can be witnessed by the number of articles published in prestigious journals such as Science and Nature (*e.g.* Viswanathan et al., 1996, 1999; Edwards et al., 2007; Sims and et al., 2008). This interest purportedly stems from observations that in foraging for food, organisms exhibit a stochastic search pattern that are not Brownian, but are rather characterized by episodic long excursions. A mathematical formalism that captures these characteristic is Lévy distributions; a general class of probability distribution functions that encompasses not only Gaussian distributions, but also so-called "long-tailed" distributions. Random walks, with run-lengths drawn from a Lévy distribution are termed Lévy walks.

The appeal of Lévy walks as a descriptor of motility is further strengthened by the observation that under certain conditions, they increase the search efficiency of the organisms that execute them. It is within this context (ie observed long-tail distributions, and predicted increased search efficiency) that Lévy walks have become ingrained in current thinking in ecological modelling, and behavioural and evolutionary ecology.

Recently however, some cracks have started to appear in the arguments for Lévy walks. Methodological and statistical re-examination of reported observations of Lévy walks in nature shows that they may not be a prevalent as previously thought (Edwards et al., 2007). Furthermore, their supposed efficiency in search behaviour is also doubtful. This latter point can be framed in two questions:

- Are the "certain conditions" under which Lévy walks increase search efficiency actually relevant to natural settings?
- Are there other, simpler, search strategies that give a even greater search efficiency?

For some reason, these questions are almost invariably ignored by protagonists of Lévy walks.

Furthermore, there is the question as to how much of observed motility patterns are determined by underlying behavioural algorithms and how much is imprinted by a non-random distribution of targets in space. In particular, resources in the marine environment are, in general, patchy – they have a clumpy distribution. It seems conceivable therefore, that simply moving in straight lines until a target is located, consuming it, and moving off again along a straight line in a random direction could reproduce long-tailed run length distributions. That is, a Levy like distribution of run lengths might be observed, but the underlying algorithm that produced it makes no use of Levy statistics at all.

Finally, there is the question of the biological mechanisms required for an organism to implement a particular behavioural algorithm. While there are any number of bio-chemical processes that can produce random events (behavioural triggers) with finite mean and variance, its is not al all certain that simple, efficient means of generating Lévy statistics are biologically feasible. That Lévy walks are elegant and fashionable mathematical entities is insufficient reason for their selection by evolutionary processes.

It is with a critical examination of these factors in mind, that I thought I would devote a few pages.

6.4.1 What is a Lévy walk

Lévy walks are random walks are composed of straight line segments with random reorientations where run length ℓ drawn from a Lévy distribution. Lévy distributions are a whole

family of probability density functions (Levy skew alpha-stable distribution) which technically are defined by the Fourier transform of their characteristic functions $\varphi(\alpha, \beta, \mu, c)$. A particular subset of this family, the symmetric, centred Levy distributions, are given by

$$\varphi = exp\left(-c|k|^{\alpha}\right) \tag{6.4.1}$$

where c is a scale factor. The probability density function is thus given by

$$p(x; \alpha, c) = \frac{1}{2\pi} \int_{\infty}^{\infty} exp\left(-ikx - c|k|^{\alpha}\right) dk \tag{6.4.2}$$

Lévy distributions per se are quite complex mathematically entities, and are hardly ever used in practical applications of modelling. A much simpler distribution, and one that captures the essential characteristics of a Lévy distribution is the Pareto distribution (also known as Zipf's law), where the cumulative probability density function is given by:

$$P\{\ell_i > \ell\} = \left(\frac{\ell}{\ell_m}\right)^{-\alpha} \tag{6.4.3}$$

This is the probability of finding a particular realization ℓ_i greater than ℓ, and where ℓ_m is a scaling factor. This is defined only for $\ell > \ell_m$. The associated probability density function

$$p(\ell) = \alpha {\ell_m}^{\alpha} \ell^{-(\alpha+1)} \tag{6.4.4}$$

so that $p(\ell)d\ell$ is the probability of ℓ lying in the interval $[\ell, \ell + d\ell]$.

Both the Lévy and Pareto distributions have the asymptotic property that for for large ℓ, the probability density function follows

$$p(\ell) \sim \ell^{-\mu} \tag{6.4.5}$$

where $\mu = \alpha + 1$. Depending on choice, either α or μ is used a the power law parameter.

Simulating a Lévy distribution: If r_i is a uniform random variable in the range $[0, 1]$, then a simple Pareto distribution can be constructed via

$$\ell_i = c r_i^{1/(1-\mu)} \tag{6.4.6}$$

which gives a power law distribution of order μ in the range $[c, \infty]$. A simple shift by c

$$\ell_i = c(r_i^{1/(1-\mu)} - 1) \tag{6.4.7}$$

gives a similar distribution over the full positive range $[0, \infty]$. The scale factor c is usually set to unity as this has no direct influence on the power law behaviour.

6.4.2 Search efficiency in a random target field

From a simple geometric agrement, we can deduce that for a random (*i.e.* Poisson) distribution of prey, the most efficient search behaviour is to simply swim in a straight line (*i.e.* ballistically). The volume swept out (that is searched) by an organism with detection distance R swimming in a straight line with speed v increases with time t as $V_b(t) = 2\pi vt$. If the organisms were to swim in any sort of zig-zag pattern with the same speed, irrespective of the statistics of the run-lengths – Brownian or Lévy or any other distribution – the volume searched would always be less than $V_b(t)$ for the simple reason that a zig-zag path will always

incur a degree of overlap. Thus, for an organism with no knowledge of the prey distribution, the most efficient search strategy is to move in a straight line.

Why then is it that so many recent articles start with a statement along the line that "Lévy walks with $\mu \approx 2$ are the optimal search strategy for scarce fixed targets that are randomly located"? Semantically it may be correct but it is very misleading. If the organisms' only goal is to find a target, then this may hold, but if it subsequently consumes the target (as is the ultimate goal of searching for food), the the optimal search strategy changes to a $\mu \to 1$ Lévy walks, that is ballistic motion. To be fair, the original article from which most of this stems, make precisely this point: For an organism with no knowledge of the target distribution, the most efficient search strategy is to (1) move ballistically ($\mu \to 1$) if the targets are destroyed, or to (2) perform a Lévy walks ($\mu \approx 2$) if targets are not destroyed.

The underlying reason why $\mu \approx 2$ Lévy walks are so efficient for non-destructive searches of random targets is that the searcher revisits the same target many times in rapid succession before setting off on a long excursion to a new target. Numerically this may give an apparent high efficiency, but in practical terms, its hardy efficient at all.

The argument usually put forward as to why the non-destructive search strategy is applicable, is that the targets represent renewable patches of resource. As an example, you can think of these now as berry bushes, so that foraging on a target does not completely deplete the food supply, and the target can be revisited at a later time once the berries have had a chance to recover. But this begs the question, why move at all? Clearly this has to do with the relative time for regeneration compared to the time required to find a new patch. A more robust model would surely follow the lines: When targets represent patches of resource, the organism should (1) move ballistically to find patches. Depending on the relative time for the patch to regenerate T_r versus the time to find an other patch T_s the organism should then (2) remain in the patch indefinitely ($T_r < T_s$), or (3) move ballistically to find a new patch ($T_r > T_s$). Simply put, this is Charnov's marginal value problem (as in optimal foraging theory, *cf.* Charnov (1976), and discussed in Sec. 7.2.2).

6.4.3 Search efficiency in a patchy target field

The distribution of targets in aquatic and marine environments is almost universally patchy. When foraging is such an environment, the forager may not have any pre-existing knowledge on the spatial locations of the targets, but there is one crucial piece of information it can glean from its experience. That is:

> Having just found a target, there is an elevated probability of another target being close by. There is benefit to be gained by an intensive search in this vicinity.

This is precisely what patchiness implies, and it is the sort of information (*i.e.* readily available, germane to fundamental life processes and easily acted on) on which evolution could engender behavioural adaptations. The corollary to this of course is:

> Having *not* just found a target, there is an elevated probability of there *not* being another target in the immediate vicinity. Efficiency dictates to move on as fast as possible, that is in s straight line.

An efficient search strategy in a patchy target distribution should be meshed to recent encounters with individual targets.

A simple search behaviour that includes this imperative is an encounter modulated random walk following a set of simple rules:

- Travel ballistically until a target is found.
- After a target is found and consumed, switch to an intensive local search.
- If the interval between the discovery of one target to the next exceeds some time, abandon the local search and switch to ballistic mode.

This is precisely the search strategy proposed by Benhamou (2007), and which invariably shows a higher search efficiency than any Lévy search.

Before closing this section, there are 3 points that should be made.

Firstly, there is a world of difference between the patterns that emerge from an organism interacting with its environment and the behavioural algorithm that produced it. In his comment on Benhamou (2007), Reynolds rather disingenuously claimed that this simply proved the point since ballistic and Brownian motion are end members of the Lévy continuum, so that a composite random walk is simply a Lévy walk in disguise. But this is not the point. A composite random walk predicated on the searcher's encounter with targets is inherently more efficient because it utilizes information on the patchy distribution. Albeit extremely basic, this information meshes the search strategy to the target distribution. Any who seek an evolutionary imperative for behavioural algorthims based on Lévy search strategies (*e.g.* Bartumeus, 2007) will find it scientifically infertile.

Secondly, foraging efficiency is not necessarily the over-riding imperative in dictating search behaviour. Specifically, a search strategy that is efficient in finding prey is often times also extremely risky. This has profound effects on the evolutionary processes shaping foraging behaviour, for while failing to efficiently find resources may be detrimental to some degree, there is nothing quite so unforgiving to an organism's reproductive success as its premature demise in the clutches of a predator. Search behaviour should therefore reflect a trade-off, between finding prey and mitigating predation risk. This is precisely a theme to be taken up in a following chapter.

The final point to make is that the debate on Lévy walks and search strategy says more about the scientific method and those engaged in its practice than it does about the natural world. To quote Karl Popperfrom *The Poverty of Historicism* (1957),

> " If we are uncritical we shall always find what we want: we shall look for and find confirmations, and we shall look away from and not see, whatever might be dangerous to our pet hypothesis"

This is a resounding injunction against what turns out to be a relatively common practice in science. Is it at all possible that the ongoing discussion of the search efficiency of Lévy walks is an example of a pet hypothesis? Far be it for me to guess at what motivates others, but I might add, that while Lévy walks may be mathematically fascinating, they probably have little relevance for behavioural and evolutionary ecology.

7 Adaptive Behaviour of Individuals

Zooplankton, despite their relative simplicity, have a wide range of behaviours – they change their swimming speed, adopt spiralling and meandering swimming paths, migrate vertically, change their feeding mode and targeted prey, have complex and varied mating behaviours, become "farmers" of bacterial cultures – some can even develop photosynthetic abilities having both autotrophic as well as heterotrophic attributes, an ability known as mixotrophy. Most of these behaviours (*e.g.* swimming speed, feeding mode, degree of mixotrophy) are plastic in that organisms can apparently adapt them to prevailing environmental conditions (*e.g.* abundance and type of prey, predator abundance, light, turbulence). This of course begs the question; why do plankton behave the way they do, and why does this behaviour change from time to time? The underlying imperative of plankton behaviour (indeed for that of all life forms) must lie in the process that produced it – namely evolution by natural selection.

This inference, that behaviour and evolution are inextricably linked, is extra information that can be mobilized to predict how planktonic organisms should behave in nature. Specifically, if we can mechanistically determine how an organism's fitness (reproductive success in its simplest interpretation) is linked to behaviour within a given ecological setting, then we can predict an organism's optimal behaviour. If we have correctly identified the main ingredients of fitness, and quantified their relationship to behaviour and environment, then the predicted optimal behaviour should correspond to the observed behaviour in nature.

This chapter deals with mechanistic descriptions linking the behaviour of individual plankton with their fitness. The argument proceeds from the observation that nearly all important life processes of plankton, indeed all organisms, are mediated by an encounter rate of one kind or another. Encounters with resources (food) sets the gross energy income of an organism, encounters with mates sets in part reproduction rates, while the mortality is in some measure the probability of encountering a predator. Fitness, in this context, can be viewed as a trade-off between beneficial and detrimental encounters.

7.1 Reproduction, survival and fitness

Evolution by natural selection is the grand unifying theory of life sciences. This of course is the core of Darwin's theory of evolution and the origin of species. The evidence for natural selection is not only to be found in the geologic record of bygone epochs; it echoes in how living organisms react and interact in the present. Specifically, there is a "rationale" behind the physical attributes of organisms, the behaviours they follow, and the life strategies they adopt, which is the direct outcome of the replication of traits from generation to generation, and their winnowing out by natural attrition. Inherited traits (here under behavioural algorithms) that confer a reproductive advantage to the individuals that express them will propagate into future generations.

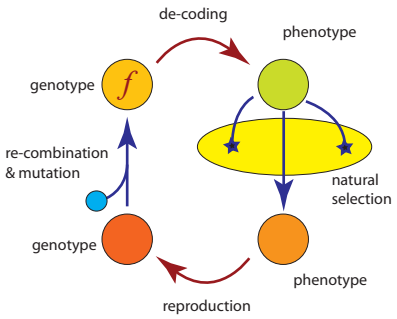


Figure 7.1: A sketch linking genotype and phenotype through natural selection and genetic recombination. Fitness is defined as the relative representation of a genotype from one generation to the next, *i.e.* counting the relative increase in genotype $f = g_{n+1}/g_n$ over successive generations. Reproductive success on the other hand simply counts surviving offspring, a measure of phenotype $r = p_{n+1}/p_n$. While these are related, they are not necessarily the same.

There is something simple and powerful in this; the combined mechanisms of Mendel's genetics and Darwin's concepts of fitness and natural selection and how they conspire to evolve and design creatures. Yet despite their simplicity, these concepts have been extremely resistent to precise mathematical expressions. This is in sharp contrast to the physical sciences where the grand laws – the laws of motion, of thermodynamics, and electro-magnetism for instance – invariably yield to elegant mathematical expressions. What could be more elegant and profound than $F = ma$, encapsulating mechanisms governing processes from the galactic to microscopic scales? It is also precise; it expresses a relationship between well-defined, *measurable* quantities; force, mass and acceleration. In life sciences, precise, practical, and measurable definitions of central components remain elusive. From Darwin's original concept, a definition of fitness seems simple enough; the expected number of offspring an individual has over its lifetime. However, with a deeper understanding of the means by which traits are inherited, it turns out that fitness is actually a genetic measure (*cf.* Fig. 7.1). To give it its modern definition, fitness is the capability of an individual of certain genotype to reproduce, and is usually equal to the proportion of the individual's genes in all the genes of the next generation. While fitness is formally a measure of ultimate genotype success, it is consequence of the proficiency of the phenotype a genotype codes for – the physical attributes and behaviour of individual organisms – and how well it meets the challenges of natural selection.

In the following I will be perhaps less than rigourous in my use of fitness, and will tend to use it synonymously with reproduction success. In some sense, I will restrict my view to phenotype, specifically how behaviour within certain environmental settings impacts reproductive success. The underlying premise is that reproductive success is a good proxy for fitness. That is, I take the optimization of reproductive success is as an evolutionarily relevant parameter, much in keeping with Darwin's original idea. This is not always the case as in kin selection for instance, but it serves as a practical solution to an otherwise clouded issue.

7.1.1 Reproductive ratio

Reproductive ratio is a fundamental measure of success in nature. It is the probable number of surviving off-spring an individual will have. In a stable population and environment, an individual with reproductive ratio 1 simply replaces itself. Individuals with reproductive ratios greater than 1, will have a greater representation in future generations (greater fitness) while those with with reproductive ratios less than 1 are on the slow road to dying out.[1] At any given time, the fittest individuals are those with the highest reproductive ratio.[2] A useful means of thinking of this is in terms of the invisibility of a "mutant" into a resident population of similar competing organisms faced with the same environment, both physical (*e.g.* space, temperature, light *etc.*) and biotic (*e.g.* nutrients, prey, predators). If the mutant invader has a reproductive value greater than the resident, it will (all things being equal) eventually come to dominate the population.

A precise definition of reproductive ratio can be got by considering a stationary population of a large number N of identical organisms. By stationary I mean that the statistical demographic properties of the population remain the same over time - *e.g.* there is no seasonal cycle, at least on the time scale of a life cycle. I assume further that statistically each individual experiences the same environment, the same resources and predators, and the same competition. Under these assumptions, the population can also the considered as N observations of the same individual in slightly different realizations of the same ecosystem. Borrowing terminology from statistical mechanics, this is a stationary ergodic process.

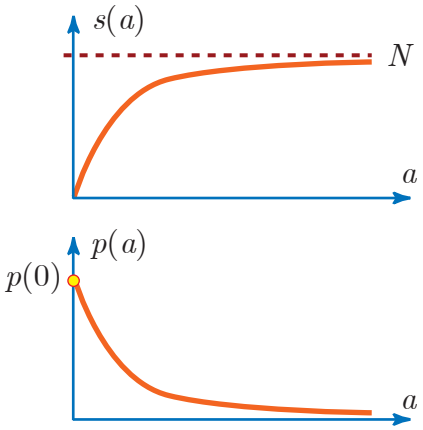


Figure 7.2: The cumulative age distribution $s(a)$ and age frequency distribution $p(a)$ for a population. Under the assumption of stationarity and ergodity, the probable survivorship of an individual can be $\phi(a)$ can be obtained directly as $\phi(a) = p(a)/p(0)$.

Much of the relevant information on individual can thus be derived from a life table of the population (*cf.* Fig. 7.2). The age frequency distribution $p(a)$ is thus given by

$$p(a) = \frac{ds}{da}\frac{1}{N} \tag{7.1.1}$$

That is, $p(a)da$ is the probability of drawing an individual from the population that is between

[1] For sexual reproduction, a reproduction ratio of 1 implies half a contribution to each of 2 offspring.

[2] This is true only in a statistical sense. Since the number of offspring depends only in part on fitness (something that may be considered deterministic) and in part on a stochastic element (the bad luck of being at the wrong place at the wrong time) as in a Galton-Watson process, it is certain that a lineage will eventually go extinct, even when the mean reproduction ratio is 1.

age a and $a + da$. Of course, $p(a)$ is normalized, so that

$$\int_0^\infty p(a)da = 1 \tag{7.1.2}$$

In a similar way to age structure, a demographic distribution of reproduction rate $\alpha(a)$ can be constructed from a cumulative reproductive rate distribution, $\sigma(a)$ by counting all the offspring $i(a)$ produced by members of the population younger than a over a period of time t. That is, $\sigma(a) = i(a)/t$, and $\alpha(a) = 1/i_{max}di/da$.

We can now define $\phi(a)$ the probability that an individual survives up to age a. It is closely related to the age frequency distribution given the stationary ergodic assumption above. Specifically, $\phi(a) = p(a)/p(0)$, the factor $p(0)$ coming directly from the constraint that the probability of an individual surviving to age 0 must be 1. An interpretation of $p(0)$ can be got from the observation

$$\int_0^\infty \phi(a)da = T \tag{7.1.3}$$

where T is the mean life expectancy of an individual. Thus $p(0) = T^{-1}$, the expected mean mortality rate of an individual.

The total expected number of off-spring an organism has within its lifetime is given by

$$r_0 = \int_0^\infty \phi(a, e, b)\alpha(a, e, b)da \tag{7.1.4}$$

This is very general. Any number of external or internal effects can influence ϕ or α at any time. Some may be age dependent (a) – newly born individuals for instance may not be capable of reproducing – some may be fluctuating in some quasi-random fashion (e) – the number of predators for instance effecting survivorship – while some depend on the "choices" (b) – behaviours and strategies – an individual makes in the face of prevailing environmental conditions.

We can get a simple definition of reproductive ratio by noting that the mean reproductive rate of the population is

$$\bar{\alpha} = \int_0^\infty p(a)\alpha(a)da = p(0)\int_0^\infty \phi(a)\alpha(a)da = \frac{r_0}{T} \tag{7.1.5}$$

Or, in other words

$$r_0 = \bar{\alpha}T \tag{7.1.6}$$

reproduction ratio of an individual is the mean reproductive rate of a population of like individuals times the mean life expectancy of an individual. Conceivably, we could have arrived at this result in a much more direct way, but taking the long way round has help shape precise definitions that are founded in measurable quantities.

We can now ask the question; having survived up to an age t, how should an organism behave under given circumstances, in order to maximize its fitness. We can break the integral into piece, a component due to past experience, and a component of future expectations. That is

$$r_0 = \int_0^t \phi\alpha da + \int_t^\infty \phi\alpha da = r(t) + r_v(t) \tag{7.1.7}$$

where $r(t)$ is the organism's past reproductive success, and $r_v(t)$ is Fisher's reproductive value – its expected future reproductive success. Clearly, at time t, an organism's strategy should be to maximize $r_v(t)$.

7.1.2 Survival and mortality rate

What is the relationship between the $\phi(t)$, probability of surviving up to age t, and the instantaneous mortality rate $\mu(t)$ experienced by individuals. Consider a cohort of initially n_o individuals. The number of surviving individuals $n(t)$ in the cohort follow the differential equation

$$\frac{dn}{dt} = -\mu(t)n \tag{7.1.8}$$

which has the general solution

$$n(t) = n_o e^{-\int_0^t \mu(s)ds} \tag{7.1.9}$$

The probability of any given individual surviving up to time t is thus

$$\phi(0,t) = e^{-\int_0^t \mu(s)ds} \tag{7.1.10}$$

It can be immediately seen from Eq. 7.1.10 that survival is a multiplicative process. That is, the probability of surviving from age 0 to $t+\Delta$ is

$$\begin{aligned} \phi(0,t+\Delta) &= e^{-\int_0^{t+\Delta} \mu(s)ds} = e^{-\int_0^t \mu(s)ds - \int_t^{t+\Delta} \mu(s)ds} \\ &= \phi(0,t)e^{-\int_t^{t+\Delta} \mu(s)ds} = \phi(0,t)\phi(t,t+\Delta) \end{aligned} \tag{7.1.11}$$

Also, if the organism goes through a non-reproductive juvenile stage up to age t_m, its life time reproductive value is given by

$$r_0 = \phi(0,t_m)\int_{t_m}^{\infty} \alpha(t)\phi(t_m,t)dt \tag{7.1.12}$$

This shows that while a juvenile, the individuals primary aim should be survival - by achieving a high growth rate to "out grow" size dependent mortality rates for instance. The best strategy for this may be different from that later in life when reproduction is the order of the day.

Returning to the reproductive value at time t:

$$r_v(t) = \int_t^{\infty} \alpha(a)\phi(t,t+a)da \tag{7.1.13}$$

If there is no additional information available to an organism on the possible outcome of alternative strategies, then its best strategy at time t is to maximize its current survivorship weighted reproduction rate. That is if the organisms can expect no change in current conditions (*i.e.* the principle of "this is as good as it gets"), then its best option from an evolutionary point of view is to maximize its reproductive value. Noting that the probability of surviving over a short interval dt is equivalent to the inverse of the instantaneous mortality rate μ;

$$\phi dt = \frac{1}{\mu} \tag{7.1.14}$$

(via fundamental definitions $\phi(t,t+\delta) = N(t+\delta)/N(t) = e^{-\mu\delta}$), the relevant fitness parameter that should be maximized to ensure the highest possible future reproductive value is

$$g(t,e,b) = \frac{\alpha(t,e,b)}{\mu(t,e,b)} \tag{7.1.15}$$

Here we have made again, explicit reference to the dependence of these parameters on t, the organisms personal history including its age, e the external environment both its biotic and abiotic components, and b the behavioural strategies the organism adopts.

This expression immediately takes risk into account. Specifically, a risky behaviour $b^\dagger$ maybe of overall benefit provided the marginal increase in reproductive rate is greater than the marginal increase in mortality. That is

$$\frac{\alpha(t,e,b+b^\dagger)}{\alpha(t,e,b)} > \frac{\mu(t,e,b+b^\dagger)}{\mu(t,e,b)} \tag{7.1.16}$$

Otherwise, the risky behaviour should be abandoned.

Finally, in many instances, future expected reproductive rates depend on the net rate at which the organism acquires energy $\dot{E}$. The faster it is acquired the more potential off-spring it can produce. More importantly, energy intake can be directly related of food ingestion rates which in turn can be related directly to encounter rates. Specifically, we can see a quantifiable mechanistic link between behaviour and reproductive rate. Specifically we can write

$$\alpha(t,e,b) = \gamma\dot{E} \tag{7.1.17}$$

where the basic assumption is that γ the energy required to produce a single offspring, is independent of environmental conditions and behaviour.

7.1.3 Reproductive value in a time varying population

The above proceeds on the assumption that the population is steady - that is it undergoes no significant increase or decrease due to overall growth or mortality. A more rigourous definition would take this into account. Specifically, how can the growth rate of a population be related to the reproduction and mortality rates of its members in an age structured, time varying population. This problem was first considered by Euler (1760), and subsequently revisited by Sharpe and Lotka (1911) and Fisher (1930).

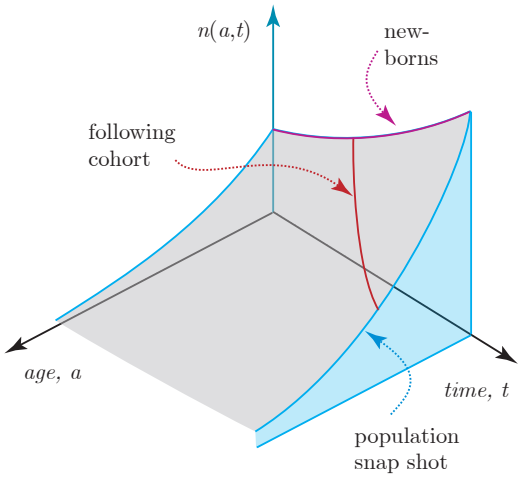


Figure 7.3: The temporal variation of the age frequency distribution in a growing population. The survivorship following a cohort has somewhat different characteristics than an instantaneous snapshot of the age distribution within the population.

Consider a population with a age frequency distribution $n(t,a)$. That is nda is the number of individuals found in the age range $[a, a+da]$ time t. Within the population, at time

t, individuals have an age dependent birth rate $\alpha(t,a)$ and mortality rate $\mu(t,a)$. The age frequency distribution thus varies in time according to:

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -\mu n \tag{7.1.18}$$

That is, the number of individuals in an age class varies over time due to inherent mortality, and the "transport" of younger to older individuals due to aging. This is the McKendrick – von Foerster equation (McKendrick, 1926; von Foerster, 1959), and its solution is a so-called renewal process. This equation is closed with the initial conditions that

$$n(0,t) = \int_0^\infty \alpha(t,s)n(t,s)ds \tag{7.1.19}$$

$$n(a,0) = n_o(a) \tag{7.1.20}$$

The first of these states that at any time t the number of individuals in the zero age class is the number of new-borns from the whole population.

A general solution to this set of equations is

$$\frac{\partial n(a,t)}{\partial t} = rn(a,t) \tag{7.1.21}$$

where r is the effective net population reproduction rate and is given by the real root of the integral equation

$$1 = \int_0^\infty e^{-ra}\alpha(a,t)\phi(a,t)da \tag{7.1.22}$$

This is sometimes referred to as the Euler–Lotka equation Fisher (1930); Charlesworth (1994)

7.2 Optimal foraging theory

Thirty years ago or so, optimal foraging theory was seen as seen as an important corner stone of ecological theory. This view has tarnished somewhat over the years, although I believe the basic premise remains sound. Optimal foraging theory is based on the observation that where ever you look, organisms seem to spend an awful lot of their time looking for and consuming food. Food of course represents an energy income, while the effort of searching, hunting, capturing, digesting and assimilating prey represent energetic costs. The central premise of optimal foraging theory is that organisms behave so as to optimize the trade-off between energetic gains and costs – specifically by maximizing their rate of net energy uptake. This appears to be a reasonable assumption as the net rate of energy income is an important aspect of an organism's survival and reproduction success. A generous energy income for instance will stave off starvation, speed sexual maturation, give a greater capacity for nurturing offspring, and allow for retreats to refuges at times of high predation risk. Optimal foraging theory addresses some aspects of the decisions an organism faces in its day to day life.

Optimal foraging theory can be broken down into three assumptions:

- The foraging strategy that is exhibited by organisms now, is that which has been favoured by natural selection in the past.

- This strategy is also that which most enhances an individual's fitness at present.
- High fitness is achieved by high net rates of energy intake.

The first two are straight forward, being restatements of Darwin's concept of evolution by natural selection as applied to foraging behaviour. The third assumption however is perhaps less than rigourous. Specifically, the link between energy intake and fitness may be more tortuous than the assumption implies. The cost-benefit balance is one aspect of behaviour-fitness link, but so is risk as we have seen in Sec. (7.1). Be that as it may, there are some valuable concepts that lend insight into optimal behaviour.

7.2.1 Prey profitability and selection

The currency of optimal foraging theory is prey profitability

$$p_i = (g_i - c_i)/t_i = e_i/(s_i + h_i) \tag{7.2.1}$$

That is the profitability p_i of food item i, is the gross gain of energy g_i the prey item represents, minus the energetic cost c_i expended in acquiring it, divided by t_i the time taken to search for and handle the prey item. This may also be written in terms of the net energy income e_i per prey and the search and handle times, s_i and h_i respectively.

Now let us suppose an organism has been feeding on a diet with mean energy content $\bar{e}$, search time $\bar{s}$ and handling time $\bar{h}$. It comes across a prey item with net energy content e_i and handling time h_i. Should the organism capture and consume this prey, or is its time better spent searching for a prey item of higher profitability? The mean rate of energy acquisition before the encounter with the new prey is

$$\dot{E}_n = \frac{\sum_j e_j}{\sum_j (s_j + h_j)} = \frac{n\bar{e}}{n(\bar{s} + \bar{h})} \tag{7.2.2}$$

If the encounter prey is included in the diet, the new mean energy acquisition rate is

$$\dot{E}_{n+1} = \frac{n\bar{e} + e_i}{(n+1)\bar{s} + n\bar{h} + h_i} \tag{7.2.3}$$

while if the prey is not included

$$\dot{E}'_{n+1} = \frac{n\bar{e}}{(n+1)\bar{s} + n\bar{h}} \tag{7.2.4}$$

Simply put, if $\dot{E}_{n+1} > \dot{E}'_{n+1}$ then the item should be consumed, otherwise the item should be passed up and the organism should continue its search. Generally, this condition of dietary inclusion is stated as

$$\frac{e_i}{h_i} > \frac{\bar{e}}{\bar{s} + \bar{h}} \tag{7.2.5}$$

the new prey item should be included if its profitability (specifically e_i/h_i) is greater than the mean rate of energy intake the organism currently experiences.

Searchers and Handlers: Some organisms spend much of their time searching for prey relative to the handling time per item: $\bar{s} \gg \bar{h}$. Eq. 7.2.5 suggests that organisms should include prey over a a wide range of profitabilities. That is searchers should be generalists. For other

organisms, handling time is the dietary bottle neck: $\bar{s} \ll \bar{h}$. The organism's overall energy intake rate is very sensitive to the profitability of new prey items. This suggests that handlers should be specialists.

Abundance of dietary components: This seems to suggest that the inclusion of a prey species depend only on its profitability and not on its concentration. That is, an organism should always reject a low profit prey item, irrespective or their abundance. This of cited result of optimal foraging theory seems to confound common sense. But read carefully. It is the only the abundance of rejected prey that is unimportant. The abundance of the included prey is actually of vital importance. The effect of abundance is masked in $\bar{s}$ and $\bar{h}$ in that all encounters contribute to the former while only encounters with included prey contribute to the latter. Going back to $\dot{E}_{n+1} > \dot{E}'_{n+1}$ it is clear that the criterion for rejection is that the mean energy intake rate if included is less that the mean energy intake rate if excluded. More importantly, if the prey item is rejected, the mean energy intake rate still decreases, but not as much.

Let us suppose an organisms is presented with a mixed diet of two species, with different profitabilities such that

$$\frac{e_1}{h_1} > \frac{e_2}{h_2} \tag{7.2.6}$$

and these are present is differing concentrations C_1 and C_2. Assuming that β, the encounter kernel for both prey species is the same, the average search time for an indeterminate prey item is $\bar{s} = (\beta(C_1 + C_2))^{-1}$. The probability that the the encounter is with a type 1 prey is $\phi_1 = C_1/(C_1 + C_2)$ and that for a type 2 prey is $\phi_2 = C_2/(C_1 + C_2)$. We can now examine two situations: the first where the organism feeds exclusively on the higher profitability prey, and the other where it includes both species in its diet in their natural ratios. The mean energy intake rate for these two diets are:

$$E_1 = \frac{\phi_1 e_1}{\bar{s} + \phi_1 h_1} \tag{7.2.7}$$

$$E_{1+2} = \frac{\phi_1 e_1 + \phi_2 e_2}{\bar{s} + \phi_1 h_1 + \phi_2 h_2} \tag{7.2.8}$$

The condition for a type 1 exclusive diet is $E_1 > E_{1+2}$, which reduces to

$$\frac{\phi_1 e_1}{\bar{s} + \phi_1 h_1} > \frac{e_2}{h_2} \tag{7.2.9}$$

That is, the condition whether a low profit prey should be included in the diet depend on the relative abundance of high profit prey. The critical relative abundance is given by:

$$\phi_1^* = \frac{s e_2}{e_1 h_2 - e_2 h_1} \tag{7.2.10}$$

For $\phi_1 > \phi_1^*$, the organism should feed exclusively on the high profit diet, but when $\phi_1 < \phi_1^*$ the organism should switch to a mixed diet.

7.2.2 The marginal value problem

An organism feeding in a patchy environment, will be confronted by the choice of either remaining within a patch and further depleting its resources, or leave the patch in search of

a fresh "fully-stocked" patch. One could think of a bird, feeding on berry bushes scattered around the landscape or a fish feeding on prey associated with coral heads.

Let us suppose that an organisms within a patch of prey, has a rate of discovery (*i.e.* encounter kernel)α and a handling time h. If the prey remains randomly distributed within the patch, the I, the rate of at which the organism consumes prey is given by a Holling's type II functional response as a function of α, h, and the within patch prey concentration $C(t)$. If the growth rate of prey is very slow compared to the grazing rate, then the prey concentration within the patch varies according to:

$$\frac{dC}{dt} = -I = -\frac{\alpha C}{1 + h\alpha C} \tag{7.2.11}$$

with initial condition $C(0) = C_0$. Despite the fact that this is a classic equation, it does not have a readily accessible solution. Parenthetically, $C(t)$ is given as the solution of

$$\ln C(t) + h\alpha C(t) = -\alpha t \tag{7.2.12}$$

although it is much more easily found by integrating the differential equation Eq. (7.2.11). Given $C(t)$, and a per capita energy content e per prey, the gross energy income while in the patch is

$$G(t) = e(C_0 - C(t)) \tag{7.2.13}$$

While in the patch, the organisms must also expend energy in maintaining its base metabolism and searching. If the energy consumption rate is m, then the net energy intake is

$$E(t) = e(C_0 - C(t)) - mt \tag{7.2.14}$$

Finally, if the energetic cost if searching between patches in E_s (this may in general be different from mT_s, the cost assuming the same metabolic cost for both within and outside the patch), then the average rate of energy acquisition per find patch – leave patch cycle is

$$\dot{E}(t') = \frac{e(C(t') - C_0) - mt' - E_s}{T_s + t'} \tag{7.2.15}$$

where t' indicates the leaving time. But how does $\dot{E}$ vary with (t')?

The simplest way to see a solution is to look at a graphical representation (Fig. 7.4). The maximum mean energy acquisition rate is that which maximizes the slope of a line from the origin to the curve $E(t)$ – this is the line that is just tangential to the curve as drawn.

7.3 Adaptive behaviour, foraging – risk tradeoff

Theoretical ecologists can be a prickly bunch. You'd be hard pushed to find a field flooded with so many die-hard dogmatists outside of a seminary. Optimal foraging theory was one topic that drew fire over the years, criticized in such memorably polemic titles as "Eight reasons why optimal foraging theory is a complete waste of time" (Pierce and Ollason, 1987), "Faith and foraging: A critique of the paradigm argument from design" (Gray, 1987), and "The meaninglessness of foraging behavior" (Ollason and Lamb, 1995). It is difficult for an outsider to see what all the fuss was about, particularly in the light of a recent review volume from which it becomes apparent that far from being a complete waste of time, (optimal) foraging theory has been a key driver in the development of behavioural and evolutionary

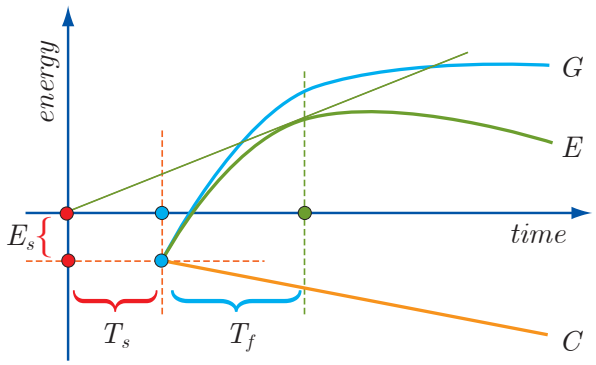


Figure 7.4: A graphical representation of Charnov's marginal value problem of foraging in a patchy environment. The net energy gain $E = G - C$ where G is the gross energy gain, and C is the metabolic cost. Here, G initially increases, flattens out and eventually decreases as metabolic costs exceed income form an ever decreasing food source. If the time to search for a new patch is T_s, and the energy required to perform this search is E_s, then the optimal time to spend within a patch is T_f, given by the intercept of a tangential line drawn from the origin to the curve $G(t)$.

ecology. To be sure, optimal foraging theory as originally introduce, was incomplete. Simply maximizing energy intake is not necessarily the path to greatest fitness. It has become apparent that foraging must reflect a trade-off between benefit and hazard – finding something to eat is important, but avoiding being eaten yourself is even more important.

One aspect of the theory that most often draws criticism is the concept of optimality. This, in a narrow view, suggest one specific behaviour that best achieves some goal function, the concept that the "optimal" in optimal foraging theory seems to underline. In a more rigourous sense though, this is a Pareto optimization problem (borrowed from economics), where multiple goal functions must be traded off against each other; optimizing one may act to the detriment of another. Within this landscape of possible behavioural options however, there are some combinations that are overall more beneficial (or less detrimental) than others. This type of optimization is entirely contingent on the trade-offs between competing goal functions.

In what follows I will try to elucidate this trade-off as it relates to the foraging behaviour of zooplankton; in terms of swimming speed and path geometry (Visser, 2007b), and in a turbulent environment (Visser et al., 2009).

7.3.1 Swimming speed - benefit, cost and risk

How fast should an organism swim? Specifically, how will changing swimming speed impact its trade-off between risks, benefits, and costs of its foraging effort? For a swimming organism such as a ciliate, copepod, or larval fish, increasing its swimming speed will impact its fitness in three ways; it will increase contact with food (positive), require greater energy expenditure (negative) and increase predation by higher contact rates with predators (negative).

Benefit: An organism, swimming at speed v through a field of immotile prey of concentration C will encounters prey at a rate $Z_c = \pi R_c^2 v C$ where R_c is the prey detection distance. Given

that the organism spends a finite time h handling prey, the rate at which it ingests prey is

$$I_c = \frac{Z_c}{1 + hZ_c} \tag{7.3.1}$$

(Holling's disk equation). If the available (*i.e.* that which can be assimilated) energy content per prey is e, then a relatively general expression for the net rate at which an organism acquires energy is

$$\dot{E} = eI_c - m - qv^2 \tag{7.3.2}$$

where m is the base metabolic and qv^2 represents the cost of swimming at speed v.

Cost: Moving trough a fluid requires effort to over come viscous drag. This is expressed in Eq. 7.3.2 for low Reynolds number swimming. Specifically, the drag on a spherical particle of size a moving a speed v is $f_{\text{drag}} = 6\pi a\eta v$, and the power required to overcome this drag is $q' = vf_{\text{drag}} = 6\pi a\eta v^2$. However, swimming is far from a 100% efficient. Stored internal energy has to be converted through various chemical path ways to propulsive power, and propulsive power has to expended both in effective forward motion as well as recovery or symmetry breaking motion. Both of these, chemical pathways and hydrodynamics, are inefficient. Estimates vary from as high as 40% to as low as 0.01%, depending on organism, definition and measuring technique. In any event, overall efficiency is expected to be low, and $\epsilon \approx 1\%$ may be taken as a reasonable estimate. The power expended $q = q'/\epsilon = 6\pi a\eta v^2/\epsilon$. Similar conversion factors can be derived for higher Reynolds number where q may have a higher power dependance on swimming speed (*e.g.* v^3), and for other shapes (ellipsoids for instance).

Risk: While swimming, an organism also increases its probability of encountering a predator. If the predator swims at a speed u, and has a detection distance to the organism R_p, then the rate at which an organism will encounter predators is

$$Z_p = \pi R_p^2 P \left(u^2 + v^2\right)^{1/2} \tag{7.3.3}$$

where P is the concentration of predators. Not all encounters with a predators will lead to the organism being captured. Many zooplankton have relatively well-honed escape abilities. Furthermore, encounters with predators are not the only causes of death. Starvation, disease, parasites and old-age also contribute to mortality. An expression for overall mortality may thus be written

$$\mu = \mu_0 + \varphi\pi R_p^2 P \left(u^2 + v^2\right)^{1/2} \tag{7.3.4}$$

where μ_0 is the background mortality, and φ is the proportion of predator encounters that lead to capture.

Fitness: Darwinian fitness - the central concept of evolutionary theory - describes the capability of an individual to increase its representation in future generations. Elements of fitness thus include growth, reproduction and survival. Precisely how these elements puzzle together to define fitness though, is not that clear, and a variety of parameterizations have been promoted in the literature (Giske et al., 1993; Kozlowski, 1996). None-the-less, a parameter that has a relatively simple interpretation and captures the gross aspects of fitness, is the ratio

$$g = \frac{\alpha}{\mu} = \gamma\frac{\dot{E}}{\mu} = \gamma g' \tag{7.3.5}$$

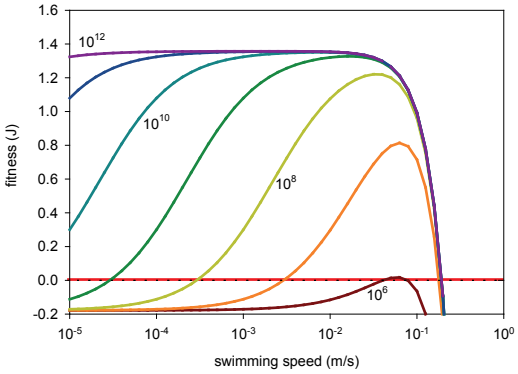


Figure 7.5: Fitness (J) as a function of swimming speed (m s^{-1}) for a variety of prey concentrations (m^{-3}). Parameters correspond to an adult copepod, *e.g. Temora longicornis* (length 700 μm; detection distance 600 μm; base metabolism $m = 2 \times 10^{-7}$ W; and handling time $h = 4$ s) feeding on a typical prey item such as *Thalassiosira weissflogii* (carbon content 260 pg C $cell^{-1}$, $e = 6 \times 10^{-6}$ J $cell^{-1}$). Background mortality is $\mu_0 = 0.1$ day^{-1}, while a concentration of visual predators *e.g.* larval herring (detection distance 6 mm), is $P = 10$ m^{-3}. It is quite remarkable that the predicted swimming speeds for adult copepods correspond relatively well in magnitude to swimming speed evident in nature.

(*cf.* Sec. 7.1). If the mean energetic cost of producing a single, surviving offspring is γ, then reproduction rate is $\alpha = \gamma \dot{E}$. Further, if the mean mortality rate is μ, then the mean life-span is $T = \mu^{-1}$, and g represents the number of surviving off-spring form this time on, all things being equal (*i.e.* reproductive value). Since γ is assumed constant, g' follows the same shape as g, and can be interpreted as the total energy acquired by the organism over its life time *in posse*.

This formulation is essentially the same as the optimal fitness criterion proposed by Gilliam and Fraser (1987) (specifically optimal behaviour is that which minimizes $\mu/\dot{E}$) and explored subsequently in theoretical studies (Gilliam, 1990; Houston et al., 1993), experiments (Gilliam and Fraser, 1987) and models (Giske et al., 1997). The underlying assumption here is that the energy required to reproduce a single, surviving off-spring is independent of an organism's current behaviour - an assumption we will follow here.

Combining the net energy income and mortality into Eq. 7.3.5 leads to

$$g' = \gamma \frac{\dot{E}}{\mu} = \frac{e \dfrac{\sigma_c C v}{1 + h \sigma_c C v} - m - q v^2}{\mu_0 + \varphi \sigma_p P \left(u^2 + v^2\right)^{1/2}} \tag{7.3.6}$$

where for brevity the interactive cross-sections are written $\sigma_{c,p} = \pi (R_{c,p})^2$. For any given external condition (*e.g.* prey concentration C, predator concentration P) optimal swimming speed is that which maximizes the function g'.

The general shape of fitness as a function of swimming speed is given in Fig. 7.5. These curves have in general have a maximum which becomes more localized (in v space) as the prey concentration decreases and swimming speed becomes constrained by energy costs.

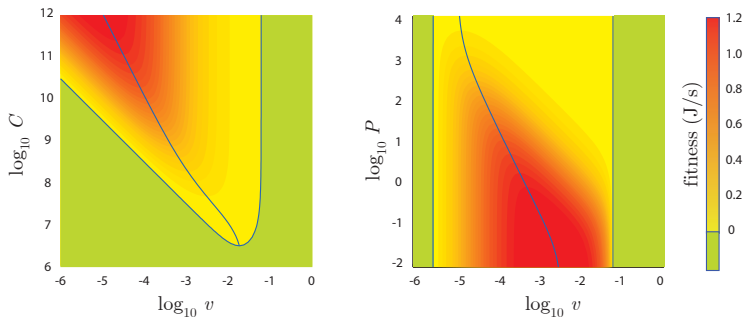


Figure 7.6: Fitness as a function of a typical adult copepod as a function of its swimming speed v(m/s), for variable prey concentration C(prey/m^3), and predator concentration P(predators/m^3).

Fig. 7.6 provides an illustration of fitness, optimal swimming speed, and how these depend on both predator and prey concentration. The parameters in this example are consistent with an adult copepod feeding on a ciliate and being preyed on in turn by a larval fish. A notable trend is that as both predator and prey concentration increase, optimal swimming speed decreases. That is:

- more prey ⇒ swim slower
- more predators ⇒ swim slower

a prediction borne out in several observations (*e.g.* van Duren and Videler, 1996). An implication of this on trophic dynamics can likewise be drawn:

- more predators *or* prey ⇒ reduced grazing pressure on prey
- more predators *or* prey ⇒ reduced trophic transfer to predators

leading to behaviourally mediated trophic cascades, an issue that will be explored later.

7.3.2 Swimming path - meanders, zigzags and spirals

Faced with the duel tasks of searching for food and avoiding predation, is there an optimal path geometry an organism can adopt? While there are many facets to this question, as a starting point I note that an organism swimming in a straight-line path presents the most efficient means of searching for prey while also exposing itself to maximum predation risk (*cf.* Sec. 6.4). An organism can mitigate its predation risk by the simple expedient of introducing a degree of curvature to its path in the form of meanders, zigzags or spirals. Executing a convoluted path, the organism is doubling back over volumes of space where it has already exposed itself to risk, and that it has evidently found "safe". Moreover, since search distance is generally smaller than the exposure distance, a moderate degree of curvature may significantly decrease exposure to risk while having a barely perceptible effect on its foraging efficiency (Fig. 7.7).

In what follows, I consider two motility patterns commonly exhibited by planktonic organisms: helical swimming and random walk motility and how these swimming paths affect the fitness of organisms that execute them.

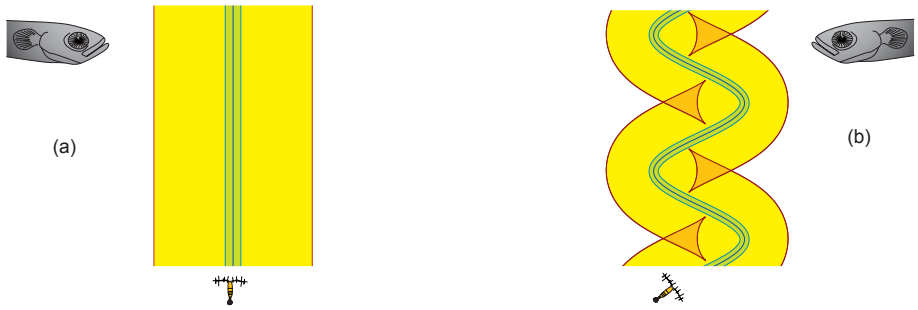


Figure 7.7: The risk-benefit of a swimming organism can be deduced from the ratio of is exposure volume to search volume. Because the distance at which an organism detects its prey is much smaller than the distance it itself is detected by it's predator, a moderate amount of curvature in its foraging path, it can reduce its risk with only negligible decrease in its search efficiency.

Helical paths: These are common amongst many plankton including bacteria (Berg, 1992), protists (Fenchel and Jonsson, 1988) and copepods (Titelman, 2001). Mechanistically, they arise due to an axial rotation combined with forward propulsion that enables organisms to "average out" asymmetries in their morphology and/or propulsive thrust, thus achieving a degree of net forward motion rather than continually swimming in circles (*e.g.* Fenchel, 2001). Organisms are apparently able to alter both the amplitude and rate of rotation (thus pitch) of their helical paths (Machemer, 2001; Fenchel, 2001). Some organisms appear to alter their rate of rotation is response to local chemical concentrations, providing an effective and efficient chemotactic ability (Crenshaw, 1996; Thar and Fenchel, 2001). In light of the above, it may also be deemed possible that helical swimming has an advantage over straight line swimming in that it potentially reduces the an organism's exposure to predation risk.

A helical path can be characterized by its amplitude A, pitch B, and speed v (Fig. 7.8). The volume swept out by a spherical absorber (radius r) following a helical path, increases linearly with time as

$$V(t) = \Omega(r, A, B)ut \tag{7.3.7}$$

where $\Omega(r, A, B)$ is the area "cut" by the spherical absorber, through a plane perpendicular to the axis of the helix, and u is the along axis swimming speed. The latter is given by $u = v \sin\theta$, where θ is the pitch angle given by $\theta = \arctan(B/2\pi A)$. The volume $V(t)$ is the volume of a helical tube of height ut. It turns out that the cut area Ω is rather tricky to calculate, and this is done numerically.

The "efficiency" of the helical path is the ratio of the new (net) volume searched or exposed to, to the gross volume. That is

$$\delta(r, A, B) = \frac{\Omega(r, A, B)}{\pi r^2} \sin\theta \tag{7.3.8}$$

where for a given path $\delta(r, A, B)$, efficiency is determined by the search or exposure radius. Efficiency varies between 0 and 1 (Fig. 7.9). When both A and B are larger than r, the volume swept out is identical to that following a straight path of length vt so that $\delta = 1$. As A and or B become small compared to r, the overlap becomes large, and the path efficiency approaches zero. When $B = 0$, the path is a closed torus, and $\delta = 0$.

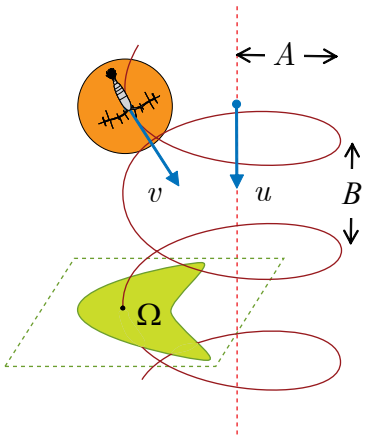


Figure 7.8: A helical swimming path is defined by amplitude A and pitch B, or pitch angle $\theta = \arctan(B/(2\pi A))$. A spherical shell, radius r and tangential speed v, whose centre travels along this path, traces out a helical tube. Ω is the cross sectional area of the tube in a plane perpendicular to the axis of the helix, and the volume of the tube increases in time as $V(t) = \Omega u t$ where $u = v \sin\theta$. An overlap in the volume traced out first occurs when $B = 2r$ or $L = 2r$ where $L^2 = (2A)^2 + (B/2)^2$.

In terms of fitness, an organism should tailor its path so as to maintain a high efficiency with respect to its search for prey ($\delta(R_c, A, B) \to 1$) while at the same time keeping its exposure to predation as low as possible ($\delta(R_p, A, B) \to 0$); R_c and R_p being the prey and predator detection distances respectively. The corresponding fitness function can be written as:

$$g(A,B) = \frac{e\sigma_c v C \delta(R_c, A, B) - m'}{\mu_0(1 + \kappa\delta(R_p, A, B))} \tag{7.3.9}$$

In this case, m' is the net metabolic cost including that due to swimming, and κ represents the increased mortality due to an increase in the number of ambush predators. Specifically

$$\kappa = \frac{\pi R_p^2 v P \varphi}{\mu_0} \tag{7.3.10}$$

While there may be some energetic costs associated with swimming along a specific helical path (*e.g.* pitch is determined by axial rotation rate which in turn requires energy expenditure to overcome viscous torque), m' is treated as being behaviourally independent. Provided the metabolic cost is small compared to the rate of energy intake ($Q = e\sigma_c C v$), the relevant functional form reduces to

$$g(A,B) = \frac{Q}{\mu_0} \frac{\delta(R, A, B)}{1 + \kappa\delta(X, A, B)} \tag{7.3.11}$$

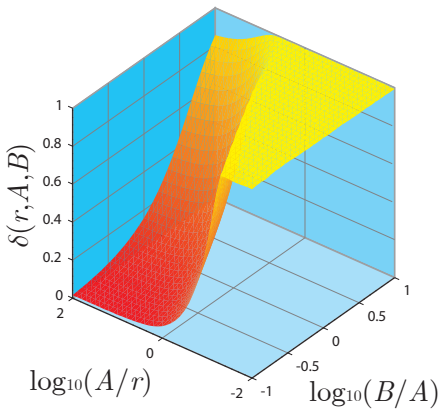


Figure 7.9: The "efficiency" of a helical path as a function of the amplitude A and pitch B of the path, and the radius of the search/exposure volume r. For large amplitude and large pitch, the efficiency approaches 1, but small values of either amplitude or pitch increase the degree of overlap so that efficiency approaches zero.

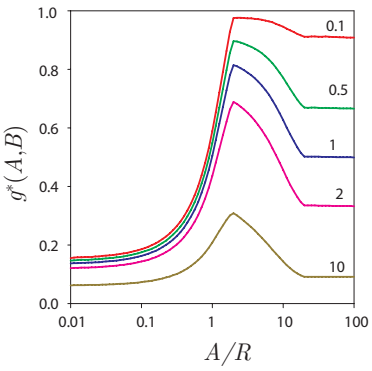


Figure 7.10: The relative fitness g^* as a function of the ratio of the helical path amplitude A and the detection distance to prey R_c. In this example, the pitch angle is set to $\theta = 0.16$ corresponding to $B = A$. The predator detection distance is set to $R_p = 10\ R_c$. Various curves are plotted for differing values of the relative predation risk κ.

Fig. 7.10 traces the relative fitness g^* given by

$$g^*(R_c, R_p, A, B) = \frac{\delta(R_c, A, B)}{1 + \kappa\delta(R_p, A, B)} \tag{7.3.12}$$

for a given fixed helical path characteristics. The ratio $R_p/R_c = 10$ is consistent with general predator - prey size ratios (Hansen et al., 1994), and the observation that detection distance often scales with body size. The corresponding fitness function is given in Fig. 7.10 for a variety of predation pressure factors. For this configuration, the optimum amplitude appears to be independent of predation pressure, and is located where the foraging efficiency first approaches 1. Geometric considerations suggest that the overlap is zero (*i.e.* efficiency is 1) when both $B/2 \geq R$ and $L/2 = (A^2 + (B/4)^2)^{1/2} \geq R$. This can be seen by considering the overlap of cross sectional search areas at half (the latter) and a full (the former) orbital periods. For the example of the path configuration considered above, $(A = B)$ the former condition is limiting so that $A^* = B^* = 2R$.

Is this a global optimum though? The organism may after all be able to vary both amplitude and pitch independently. Firstly it can be noted that the path efficiency with respect to the predator, (detection distance R_p) should be as low as possible, and lie in the lower left hand corner of Fig. 7.9. That is, $A/R_p \leq 1$ and $B/A \leq 1$. On the other hand, the foraging efficiency should be as close to 1 as possible; a condition met for $B^* \approx 2R_c$. This can be seen in Fig. 7.9 where the line $B = 2R_c$ is plotted and is coincident with the contour $\delta = 1$ in the region where $B/A \leq 1$. Combining these conditions we thus have the condition

$$R_p \geq A^* \geq B^* \approx 2R_c. \tag{7.3.13}$$

This suggests that for helical swimming, optimal geometry should conform to a spiral with a shallow pitch; i.e. the amplitude should be of the same order or larger than the pitch.

While there are no specific observations of helical swimming under the dual pressures of foraging efficiency and predation risk, some corroboration can be got from observations of nauplii swimming patterns. Titelman and Kiørboe (2003b) show typical swimming tracks of late stage *Centrophages typicus* and *Acartia tonsa* with pitch $B = 300$ to 500 μmand amplitude $A = 250$ to 500 . That is $A \approx B$. Given that these organisms have an equivalent spherical diameter $a \approx 100$ μm, their detection radius to their prey is a few times their body radius, *i.e.* $R_c \approx 200$ μm, it appears that $B \approx 2R_c$, which is consistent with the condition stated above.

Random walk path: In chapter 2, we considered the volume swept out by a spherical

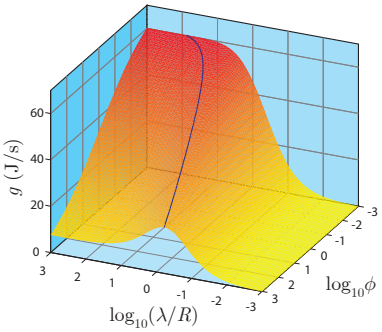


Figure 7.11: The fitness of a random walk prey search as a function of the mean free path λ, search radius R and relative predator density ϕ.

absorber execution a random walk – the so-called Wiener sausage. Given the radius of the absorber as r, and the mean free path of the random walk as λ, the rate of increase of the Wiener sausage can be approximated by

$$\beta(\lambda, r) = \frac{4\lambda}{4\lambda + 3r}\pi r^2 v \tag{7.3.14}$$

where v is the speed at which the absorber moves. For a relatively long mean free path, $\lambda \gg r$, the scanning rate (volume / unit time) becomes that of a ballistic path $\beta \to \pi r^2 v$. For a relatively short mean free path, $\lambda \ll r$, the scanning rate becomes diffusive $\beta \to 4/3\pi\lambda r v$.

The fitness of an organism in terms of its motility length scale can thus be written as

$$g(\lambda) = \frac{eC\beta(\lambda, R_c) - m}{\mu_0 + \varphi P \beta(\lambda, R_p)} \tag{7.3.15}$$

As with the previous section, if the metabolic rate is assumed small compared to Q the rate of energy intake, this reduces to

$$g(\lambda) = \frac{Q}{\mu_0}\frac{\gamma(\lambda, R_c)}{1 + \kappa\gamma(\lambda, R_p)} \tag{7.3.16}$$

where $\gamma(\lambda, r) = 4\lambda/(4\lambda + 3r)$ is the path efficiency. Fig. 7.11 traces the relative fitness for foraging ($r = R_c$) and predation ($r = R_p = 10R_c$) as a function of varying motility length scales. These are plotted for different predation risk factors κ, representing the relative risk due to encounters with ambush predators, to general background (behaviour independent) mortality risk. Essentially κ can be though of as scaling with the abundance of ambush predators. These plots show a marked maximum, particularly at intermediate values of κ. The maximum shifts towards smaller motility length scales as the behaviourally dependent predation risk increases.

The optimal motility length scale for an organism executing a random walk lies between its detection distance to its prey, and the distance at which it in turn is detected by its predators. For high risk of ambush predation ($\kappa \approx 10$), λ^* tends to the prey detection radius R_c, while at low risk ($\kappa \approx 0.3$), λ^* tends to the predator detection radius R_p. This is consistent with results found by Visser and Kiørboe (2006), which showed, for a range of planktonic organisms (bacteria, flagellates, ciliates, nauplii and adult copepods), a distinct relationship between the motility length scale and the size of the organisms itself. Specifically, $\lambda \approx 15a$ where a is the equivalent spherical radius of the organism (Fig. 7.12). Typically, the prey detection distance is a few times the radial size of the organism (*e.g.* Lenz and Yen 1993), so

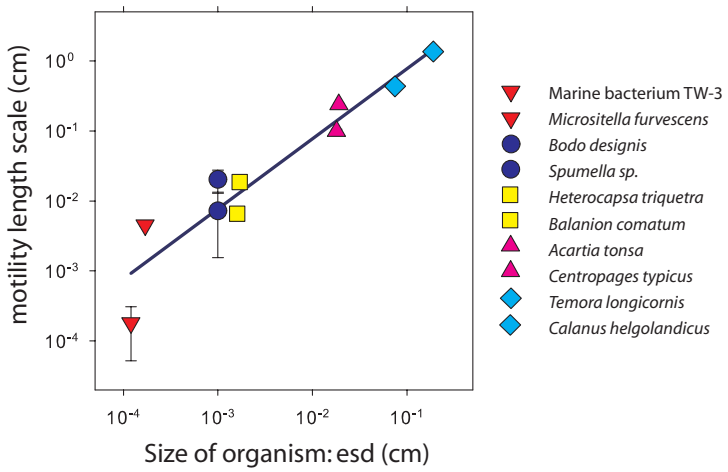


Figure 7.12: Relationship between the size of an organism (d, equivalent spherical diameter) and the correlation length scale of its swimming path (λ). Data are for those organisms analyzed in . Regression analysis suggests the relationship $\lambda = 7.67 d^{1.00} (r^2 = 0.90)$.

that $\lambda \approx 5R_c$, roughly the optimum found for an ambush predation risk of $\kappa = 1$, *i.e.* where the contribution to net mortality rates due to behaviour is the same as which is independent of behaviour.

7.3.3 Swimming in turbulence

Twenty years ago Rothschild and Osborn (1988) introduced the seminal concept to plankton ecology – that increasing turbulence increases the rate of contacts between planktonic organisms and their prey (*cf.* Sec. 4.5). In the intervening years, refinements to this concept have included:

- Turbulence increases contact rates (Marras et al., 1990; Costello et al., 1990; Mann et al., 2005) but this does not necessarily translate into increased ingestion rates. Ingestion rates only increase when prey concentrations are less than saturating (Saiz et al., 1992).
- Turbulence can interfere with an organism's ability to remotely detect (Saiz and Alcaraz, 1992; Saiz and Kiørboe, 1995) and/or capture prey (MacKenzie et al., 1994). In general, at low levels of turbulence, ingestion rate increases with turbulent intensity, but flattens off and eventually decreases as turbulence levels increase further, giving a dome-shaped relationship (MacKenzie and Kiørboe, 2000).
- Zooplankton also exhibit changes in behaviour in response to turbulence. Their metabolic rates increase as turbulence increases (Alcaraz et al., 1988). Swimming behaviour, both in speed and path geometry vary with different levels of turbulence (Saiz and Alcaraz, 1992; Seuront et al., 2004) as does their feeding mode and subsequent prey selectivity (Kiørboe et al., 1996; Caparroy et al., 1998).
- Turbulence is a highly variable environmental parameter; changing in both time and space over many orders of magnitude. In particular, vertical variations may be many orders of magnitude in the space of a few metres. Many pelagic species of copepod appear to migrate deeper into the water column to avoid high levels of turbulence in the surface (Visser, 2001; Incze et al., 2001).

These observations pose the questions, how fast a zooplankter should swim – if at all – and where should it position itself in the vertical to maximize its fitness given certain environmental conditions. The primary focus here is the effects of turbulence in that it is highly variable in both time and space, and it directly impacts fitness both positively and negatively through a number of mechanisms. In particular, turbulence enhances an organism's contacts with prey, but also its contacts with predators. A fundamental trade-off is the net rate of energy gain (energy ingested minus energy expended) versus mortality rate of a test organism following a fixed behavioural option in a given turbulent environment. The test organism can be thought of as either a member of a single species with plastic behaviour, or a generic member of a range of species that face similar opportunities, hazards and constraints. As defined above (*cf.* Sec. 7.1.2), I take the ratio of net rate of energy gain to mortality rate as a measure of fitness. As a specific example, I will consider a typical adult copepod (*e.g. Acrtia tonsa*, *Temora longicornis*, *Oithona similis*, *Centropages typicus*) commonly found in temperate shelf seas. Since feeding mode is intimately connected to motility, I consider 3 feeding modes exhibited by pelagic copepods:

- suspension feeding (*i.e.* generating a feeding current),
- ambush feeding (*i.e.* remaining motionless while relying on prey motility to bring about contacts) and
- cruise feeding (*i.e.* actively searching for prey).

Thus, deducing how optimal swimming behaviour varies as a function of depth, also says something about how feeding modes, and the species that practice them, should be arranged in space.

Ambush and cruise feeding: cost–benefit: The rate at which a organism, swimming at speed v, encounters prey in a turbulent fluid can be written as

$$Z_c = \pi R_c^2 C \left(v^2 + u_c^2 + 2w_c^2\right)^{1/2} \tag{7.3.17}$$

Here R_c is the distance at which the organism can perceive a prey, C is the prey concentration, u_c is the swimming speed of the prey, and w_c is the turbulent velocity appropriate for the organism's interaction with its prey – the latter being given by

$$w_c = \alpha_R \left(\varepsilon R_c\right)^{1/3} \tag{7.3.18}$$

where ε is turbulent dissipation rate and α is Richardson's constant. Combining these expressions with

$$\dot{E} = \frac{e\phi_c\sigma_c}{\sigma_c h + \left(v^2 + u_c^2 + 2\alpha(\varepsilon R_c)^{2/3}\right)^{1/2}} - m - qv^2 \tag{7.3.19}$$

Here metabolic cost m, handling time h, capture efficiency ϕ_c and hydrodynamic cost factor q take on the sane meaning as in section (7.3.1). The above formulation covers both ambush and cruise feeding modes, the former being for $v = 0$, and the later for $v > 0$.

Suspension feeding: cost–benefit: The cruise-ambush formulation proceeds under the tacit assumption that the organism is neutrally buoyant, and that all the swimming energy expended is converted into forward movement through the water. Many marine organisms

however, are not neutrally buoyant, and have excess densities, $\Delta\rho$, with respect to seawater of 30 to 50 kg m^{-3}, resulting in a sinking speed (assuming Stokes' regime) of

$$v_0 = \frac{2}{9}\frac{ga^2\Delta\rho}{\eta\rho} \tag{7.3.20}$$

Thus, in order to remain at a given depth, the organisms must constantly expend energy at the average rate qv_0^2. There is however, an advantage to be gained from being negatively buoyant. For the same effort, a gravitationally tethered organism (i.e. one that swims upwards to compensate for its sinking), has a 1.5 times higher clearance rate than a cruising, neutrally buoyant organism of the same size (Tiselius and Jonsson 1990). In this case, the encounter rate with prey $Z_c = 1.5R^2 v_{\text{pump}}$, where $v_{\text{pump}} = v_0$. Furthermore, a gravitationally tethered suspension feeder reduces its marginal predation risk as it contributes no relative motion to the predator encounter rate. Effort in excess of that required to hold the organism suspended will result in its forward motion - switching from suspension feeding to cruise feeding. Combining these effects we can write the net encounter rate as

$$Z_c = \sigma_c C\left(1.5v_0 + \left(v^2 + u_c^2 + 2w_c^2\right)^{1/2}\right) \tag{7.3.21}$$

The cost of swimming and maintaining a feeding current can be written as $q(v_0 + v)^2$. Here, v is the speed of the organism relative to the fluid.

This formulation (Eq. 7.3.21) merges with the ambush-cruise formulation given above (Eq. 7.3.17), as a function of the density contrast between the organism and the fluid. Specifically, when $\Delta\rho = 0$, $v_0 = 0$, and both the encounter rate and energetic cost become the same as for an ambush-cruise feeding organism.

Predation risk: The mortality rate of a zooplanker is in part due to background attrition due to parasites, disease and old-age, and in part due to direct predation due to risky behaviour tempered by environmental conditions. Combining these, we can write the overall mortality rate as

$$\mu = \mu_0 + \phi_p Z_p \tag{7.3.22}$$

where μ_0 is the background mortality rate, Z_p is the encounter rate with predators, and ϕ_p is percentage of encounters that lead to captures. As with prey encounters above, encounter rate with predators can be written as

$$Z_p = \pi R_p^2 P\left(v^2 + u_p^2 + 2w_p^2\right)^{1/2} \tag{7.3.23}$$

where P is the concentration of predators and R_p is the distance at which the organism is perceived by a predator. The latter also enters into the appropriate turbulent velocity estimate as

$$w_p = \alpha(\varepsilon R_p)^{1/3} \tag{7.3.24}$$

Detection mode: We note here that the detection distance R_p may be a function of environmental conditions. Specifically, for a visual predator, R_p depends on light and thus on time of day and depth below the surface. When we come to examine the vertical distribution of optimal behaviour, we will simulate this as a simple linear relationship

$$R_p(z) = R_{p0}L(z)/L_0. \tag{7.3.25}$$

where $L(z)$ is the light level (a function of depth, z), and R_{p0} is the detection distance at a reference light level L_0.

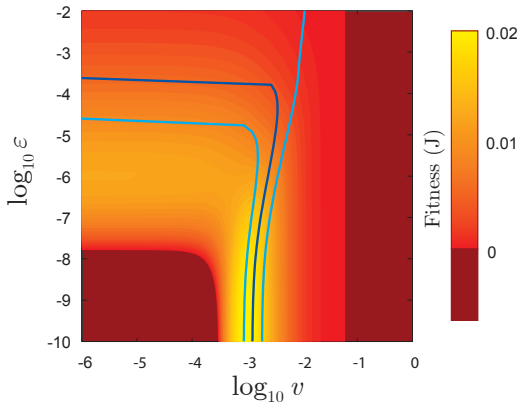


Figure 7.13: itness (in units joules: J) as a function of swimming speed v (m/s) and turbulent dissipation rate ε (m^2 s^{-3}). For a given turbulence level, the optimal swimming speed v^* is plotted, as are the values of v corresponding to $(\pm 0.95 g(v^*, \varepsilon))$. That is, an organism selecting a swimming speed within this range will at most suffer a 5% fitness penalty.

7.3.4 The fitness landscape

Formally, we will explore the fitness parameter

$$g(v, \varepsilon, C, P) = \frac{\dot{E}}{\mu} = \frac{\phi_c \dfrac{e Z_c(v, \varepsilon, C)}{1 + h Z_c(v, \varepsilon, C)} - m - q v^2}{\mu_0 + \phi_p Z_p(v, \varepsilon, P)} \tag{7.3.26}$$

It has units of energy, and in a very simple interpretation, is the net future energy acquired by an organism over its expected life-span following a given behaviour.

The resulting relationship is illustrated in in Fig. 7.13, which shows the fitness (in terms of energy acquired in an expected lifetime) as a function of swimming speed and turbulence. This example is for cruise-ambush feeding ($\Delta\rho = 0$) for a prey concentration of $P = 10^8$ m^{-3} and a predator density $Q = 10$ m^{-3}. This illustrative example is for a generic adult copepod (prosome length 700μm) with detection distance R_p =600μm, feeding on a typical prey such as *Thalassiosira weissflogii* or *Oxyrrhis marina* (size = 20μm, energy content $e = 6 \times 10^{-6}$ J cell^{-1}) while being preyed upon by a typical predator such as larval herring (MacKenzie and Kiørboe, 2000) or chaetognaths (Saito and Kiorboe, 2001): detection distance $R_q = 6$ mm.

There are 3 regions where the fitness goes to zero. When both swimming speed and turbulence are low, encounter rates with prey are too low to meet base metabolic costs. When swimming speeds are high, the cost of swimming exceeds any energy income from foraging. Finally, when turbulence is high, ingestion of prey becomes saturated while predation risk increases so that fitness decreases. Bracketed by these zones, fitness exhibits a dome-shaped relationship, both with respect to swimming speed and turbulence.

Fitness of course, is not just a function of swimming speed and turbulence, but a whole range of external conditions to which an organisms can adapt its behaviour. Even in the simple model considered here, 4 additional parameters have a direct impact: prey abundance, prey motility, predator abundance and predator motility.

7.3.5 Mitigation by migration

Within the context of fitness, two parameters, light and turbulence have an inherent vertical structure. These, together with nutrients stamp a vertical structure on the profile of primary production. Thus important components of the fitness trade-off have a vertical structure; prey abundance, predation risk, and turbulence. Given that an organisms can migrate vertically, where should it position itself, and/or what feeding mode should it adopt to optimize this trade-off.

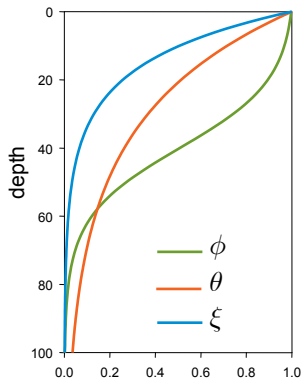


Figure 7.14: Vertical structure functions for prey abundance ϕ, light θ and turbulence ξ.

As a crude representation of the vertical structure, consider 3 profile functions, $\phi(z), \theta(z), \xi(z)$ that determine the vertical distribution of prey, the attenuation of light with depth, and the vertical profile of turbulent dissipation rate respectively. That is

$$P(z) = P_0\phi(z) \quad ; \quad \phi(z) = \frac{1}{2}\left(1 - \tanh\left(\frac{z - z_{P0}}{z}\right)\right) \tag{7.3.27}$$

$$L(z) = L_0\theta(z) \quad ; \quad \theta(z) = \exp\left(\frac{-z}{z_{L1}}\right) \tag{7.3.28}$$

$$\varepsilon(z) = \varepsilon_0\xi(z) \quad ; \quad \xi(z) = \exp\left(\exp\left(\frac{-z}{z_{\varepsilon 1}}\right) - 1\right) \tag{7.3.29}$$

where z is the depth below the surface. The prey distribution approaches P_0 towards the surface at depths shallower than z_{P0}, and exponentially approaches zero below this depth. The thickness of the transition zone is controlled by z_{P1}. Light follows a Lambert-Beer law with extinction coefficient $1/z_{L1}$. For turbulence, we essentially use an exponentially decreasing function for $log10(\varepsilon)$, with e-folding scale $z_{\varepsilon 1} = 15$ m (MacKenzie and Leggett 1991).

At any given depth z, prey concentration, light intensity (and thus predator perception distance via equation (10)), and turbulent dissipation rate can be specified. The optimal swimming speed and maximum attainable fitness can thus be estimated via equation (12) as functions of depth. The depth at which maximum attainable fitness peaks is the optimal depth where, *ceteris paribus*, an organism should try to position itself.

In general, it appears that organisms should seek out low levels of turbulence, and position themselves at a depth where the marginal decrease in prey abundance and ingestion with respect to depth exactly balances the marginal decrease in predation risk (*e.g.* due to decreasing turbulence with depth, decreasing light or both). In general, organisms that can, should migrate down with increasing surface turbulence, a process governed in this case by

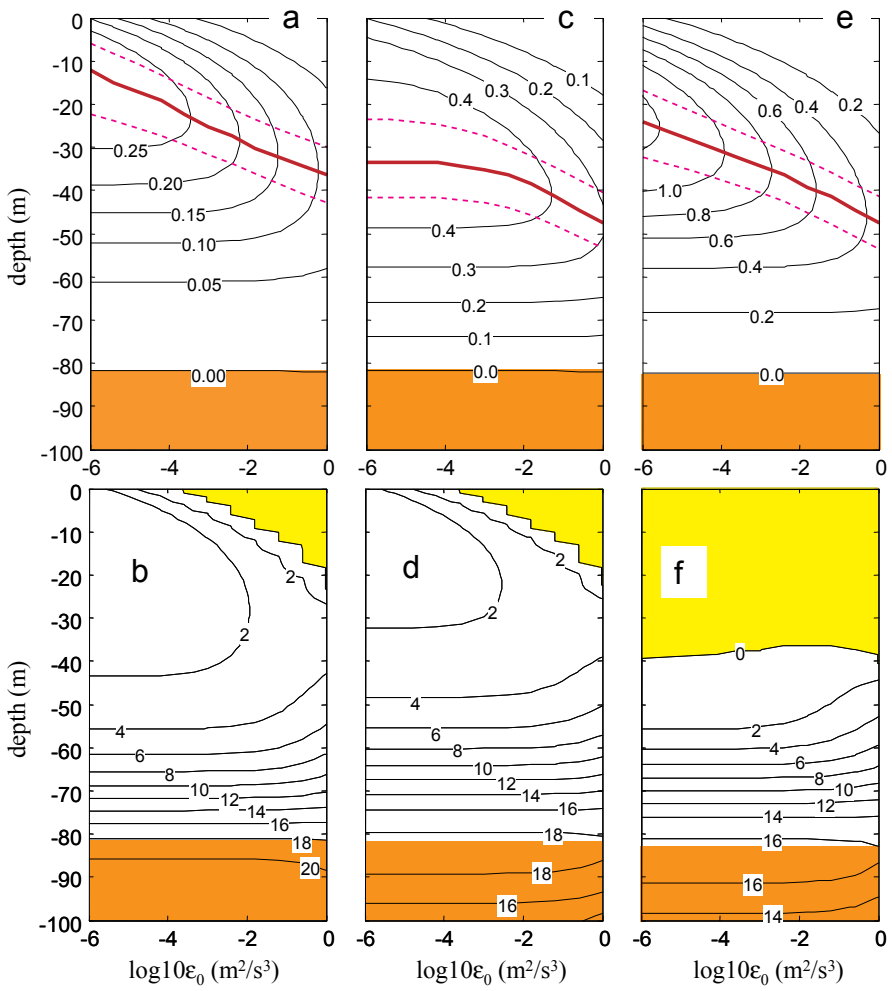


Figure 7.15: The maximum fitness g (J) (A,C,E) and optimal swimming speed v (mm/s) (B,D,F) as functions of depth and surface turbulent dissipation rate ε_0 (m2/s3) for (A,B) a neutrally buoyant organism under non-visual predation risk, (C,D) a neutrally buoyant organism under visual predation risk and (E,D) a negatively buoyant organism under visual predation risk. Depth of maximum attainable fitness is indicated by the solid red line, with the dotted lines demarking the 95% level. Non-viable regions ($g < 0$) are shade orange, while yellow indicates regions where ambush feeding (B,D) and suspension feeding (F) are optimal.

fitness considerations rather than impaired foraging ability as is usually argued (Visser et al., 2001). There is a fairly good agreement with observed turbulence related downward migration of *Oithona similis* (Visser et al. 2001), a cruise-ambush feeder, which, at night appears to migrate downwards 2.5 m per decade of surface turbulent dissipation rate, consistent with results (Fig. 7.15a). Whether this optimal depth-seeking behaviour plays out in nature, depends of course on the ability of organisms to hold their position in the face of turbulent stirring. Large organisms such as adult copepods appear to have this ability, while this is less so for nauplii and even less for ciliates (Maar et al., 2003).

7.4 Diffusion, search and time budgets

Ambush predators such as nauplii of marine copepods, (Buskey et al., 1993; Titelman, 2001) rely on the motility of their prey to bring about encounters. If the prey's motility conforms to a random walk, then the underlying encounter process is diffusive in nature (provided run length to encounter distance is small (Visser and Kiørboe, 2006)). Starting from an initially undisturbed prey field, the organism's encounter rate with prey is initially high but falls off with time as the local prey field becomes depleted. That is, the encounter rate becomes diffusion limited. To alleviate this, the organism can move to a new, undisturbed part of the prey field where its encounter rate is again high (Titelman and Kiørboe, 2003b). Each translocation, however, will expose the organism to some elevated risk of predation (Titelman and Kiørboe, 2003a). Functionally, this is very similar to the classic problem of optimal foraging on patches (Charnov, 1976) when predation risk is taken into account (Gilliam, 1990).

The most accessible approach to include time budgets and time dependent energy intake into the analysis, is to the a more general formulation of the fitness parameter:

$$g(b,t) = \int_0^t \phi(b,s)\dot{E}(b,s)ds \tag{7.4.1}$$

where, $\phi(b,t)$ is the probability of an organisms surviving to time t while executing behaviour b, and the integral is the net energy acquired in the same interval (*cf.* Sec. 7.1). This is the basis of the analyses introduced by Mangel and Clark (1988), and it can be noted that for the case where the rate of energy acquisition and survival are time independent, and taking into account that the appropriate integration time scale is the expected life span of the organisms (*i.e.* μ^{-1}), this expression reduces (to within a constant factor) to the fitness parameter in Eq. 7.1.15 and Eq. 7.3.5.

For an ambush predator, its encounter rate with motile prey is given by

$$Z(t) = 4\pi DRC\left(1 + \frac{R}{\sqrt{\pi Dt}}\right) \tag{7.4.2}$$

where $D = w\lambda/3$ (e.g. Berg (1992)) is the diffusivity of its randomly moving prey, which swim at speed w and with a motility length scale λ. The total energy acquired follows

$$\Sigma(t) = e\int_0^t Z(t)dt = 4e\pi DRCt\left(1 + \frac{2R}{\sqrt{\pi Dt}}\right) \tag{7.4.3}$$

which is initially rapid ($t < R^2/D$) but falls off with time. If this were the only consideration, then an optimal foraging time could be derived in a manner directly analogous to Charnov's marginal value problem (*cf.* Sec. 7.2.2). However, in addition to energetic costs in moving from place to place, predation risk may differ when swimming and when feeding. The fitness for a single jump-feeding bout cycle of duration $t = f + s$ is given by

$$g_1 = \left(\Sigma(f) - \left(mt + \frac{qd^2}{s}\right)\right)\exp(-\mu_s s - \mu_f f) \tag{7.4.4}$$

where s is the time spent swimming, f is the time spent feeding, and μ_s and μ_f are the mortality rates while swimming and feeding respectively. The distance d is the length scale of the depletion zone beyond which the organism should move. However, fitness per cycle (or

for any fixed time interval) is not the same as life-time fitness. With regard to the latter, the question of how many cycles can be accommodated in an expected life span also becomes an issue. In particular, the fitness after n cycles is

$$g_n = n\left(\Sigma(f) - \left(mt + \frac{qd^2}{s}\right)\right)\exp(-n\mu_s s - n\mu_f f) \tag{7.4.5}$$

from which it can be deduced that the number of cycles in an expected life span is

$$n = \frac{1}{\mu_s s + \mu_f f} \tag{7.4.6}$$

Thus, to within a constant, the life-time fitness is given by

$$g(f,s) = \frac{\Sigma(f) - (mt + qd^2/s)}{\mu_s s + \mu_f f} \tag{7.4.7}$$

(consistent with that derived by Gilliam (1990)). Finally, predation while swimming depends on swimming speed, so that $\mu_s \propto v = d/s$. Thus, swimming fast will incur a greater mortality rate but it will be sustained for a shorter period so that risk, $\mu_s s$ is independent of swimming duration. Eq. (7.4.7) thus becomes

$$g(f,s) = \frac{\Sigma(f) - (mt + qd^2/s)}{\mu_0(f+p)} \tag{7.4.8}$$

where μ_0 is the mortality rate while feeding and $\mu_0 p$ is the mortality associated with traversing the distance d. That is the mortality associated with traversing the distance d is equivalent with feeding for p seconds.

Solving Eq. (7.4.8) for $\partial g/\partial s = 0$ and $\partial g/\partial f = 0$ gives the optimal swimming and feeding durations as being

$$s^* = d\sqrt{q/m} \tag{7.4.9}$$

$$f^* = \frac{1}{\xi^2}\left(p(1-m') - k + ((p(1-m') - k)^2 + p\xi^2)^{1/2}\right)^2 \tag{7.4.10}$$

where $k = 2d(qm)^{1/2}$, $\xi = R(\pi D)^{-1/2}$ and $m' = m/(4e\pi RDC)$. The optimal swimming duration is independent of either prey concentration or predation risk.

Nauplii often exhibit jump-ambush behaviour with jump frequencies ranging from 1 min^{-1} to 3 s^{-1} (Titelman, 2001). With parameters corresponding to naplii (Visser, 2007b), the optimal swimming (jump) duration calculated from Eq. 7.4.10 for this example is $s^* = 0.14$ s, and is independent of predation risk and prey abundance. At high prey concentration, optimal feeding duration f^* increases markedly with predation risk - the benefit of moving beyond the diffusive boundary layer becomes negligible compared to the risks involved. At intermediate values of the predation risk factor, and for a broad range of cell concentrations, the optimal feeding duration for these parameters range between 0.1 and 10 s. These values of f^*, and that for s^* compare favourably with observed values (Titelman, 2001).

7.5 Mechanisms, fitness and behaviour

Given the general development of evolutionary and behavioural ecology through the 20^{th} century, I find it somewhat odd that so few of their underlying concepts have made their way

into marine ecology. By-and-large the modelling of marine ecosystems is still dominated by a physico-chemical view – the ecosystem as a stirred chemostat with life as a chemical reaction – for which evolution is deemed irrelevant. When such models fail to mesh with observations as the invariably do, modelers usually point to poorly defined rates, the provenance of empiricists, while empiricists point to the inability of models to capture the behaviour of organisms.

But there is a wealth of extra information that can be applied to the problem. Most importantly, organisms do what they do for a reason. This is the light of evolution, at least as it pertains to behavioural ecology. If the reason can be deduced mechanistically, then the how, when and where of behaviour can be understood and predicted.

Here are a few concepts that may prove useful.

Functional Response Curves: Ecosystems are generally shaped by the rates by which organisms interact with each other and their resources. These rates and how they vary with resource abundance, take on a near universal form: they increase linearly at low abundance, and then saturate at higher abundance. This form is true in planktonic systems from the uptake kinetics of dissolved matter, to the photo-response of phytoplankton, to the ingestion rate of zooplankton. The form of these functional responses is governed by two aspects: collecting and processing (Fig. 7.16). Each of these has a cost associated with it; finding prey for instance incurs an energy cost in searching while processing prey requires maintenance of biochemical pathways to breakdown and assimilate ingested prey. The same principles apply to light-harvesting or osmotrophy. That is, there is an inherent trade-off between energy income from gathered and processed resources, and the relative cost of the gathering and processing mechanisms.

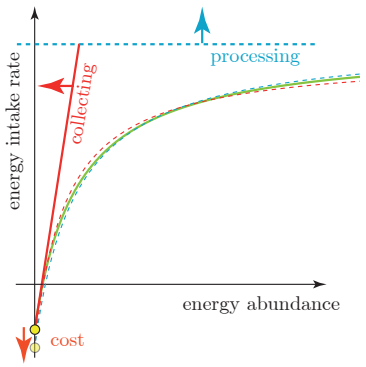


Figure 7.16: Consider an organism collecting and processing resources (energy) according to a classic Holling's II functional response. How should it invest any surplus energy? Into collecting or processing apparatus? Clearly this will depend on the cost of investment in one or other element, and the benefits these investments will accrue given the current level of resource abundance.

Sexual Selection: Different pelagic copepod species use different mechanisms for mate location, some use pheromone trails, or pheromone clouds while others use hydromechanical signals. The fitness of a male is its life-time encounter with receptive females; encounters that depend on the concentration of females, the distance at which males can perceive them (long for a pheromone trail, comparatively shorter for hydromechanical signals), and the search strategy of males (*e.g.* swimming speed and pattern). In general, the predation risk for a female, for example, while slowly swimming and laying out a highly specific chemical trail, is lower than for a male searching for the trail. The effective sex ratio is thus, in general, skewed towards females in a way inversely proportional to the relative mortality of the two sexes (Kiørboe, 2006). That is, sex ratio of the population is an emergent property of the mate finding strategies of individuals, and the biased sex-ratio may occasionally be an im-

portant element in regulating population dynamics. Understanding this system involves a mechanical model of encounter rates with mates and predators as a function of behavioral mode (activity), and a model of the reproductive success of the behavior. This success may depend strongly on the particular mating system of the species, for instance, if it involves a spermatophore (or many) or not. If the female accepts only one spermatophore, then one would expect males to pay lower attention to predation risk and a higher skew in adult sex-ratios than in systems with continuous mating. The point here is that modelling of such questions requires a deep integration of mechanical and evolutionary reasoning.

Personality and risk aversion: Not all members of a populations behave the same, neither does the same individual behave the same every time it encounters the same conditions. In a human population we would term this personality. An absolute maximization of fitness may be sub-optimal in that it locks organisms into a local, rather than a perhaps more global fitness maximum. There is also the problem of cognition in that behaviour can only be attuned to environmental factors that can be sensed. In the context here, mortality rate is perhaps most important - an organism may have a good sense of how many prey are available through its recent feeding success, but a direct measure of predation pressure is a rather dead-end pursuit. Proxy signals such as light, chemical smells or hydrodynamic regime may be of some use, and it is perhaps with regard to the more poorly known fitness elements that behavioural personalities are expressed - risk-aversion for instance (Fiksen et al., 2007).

8 Adaptive Dynamics in Populations

Ecological communities are structured through direct and indirect interactions active in food-webs (Pimm and Lawton, 1977, 1978). Interactions such as predation, mutualism, competition, commensalism and parasitism are all modes that inter-connect various parts of a community, shape its topology and to a degree, govern overall ecosystem dynamics. Moreover, there is a growing realization that there are subtler, behaviourally-mediated, indirect effects in play (Dill et al., 2003; Schnitzer, 1993). For instance, zooplankton migrate vertically avoiding sunlit surface waters during day-light hours giving their phytoplankton prey a temporary respite, piscovous fish patrol the periphery of zooplankton swarms giving protection to zooplankton while using them as "bait", whales and tuna drive fish upward to create otherwise unavailable foraging opportunities for marine birds, and fishermen follow birds to favourable fishing grounds.

In terms of ecosystem dynamics, such behaviourally-mediated effects are perhaps as important in determining the ever changing interaction strengths (and thus coupling) between populations as are the ups and downs of their relative abundance.

Whatever the architecture of ecological communities, the interactions that control their shape, are ultimately enacted at the level of individuals whose behaviour is attuned solely to their evolutionary self-interest. Adaptive foraging behaviour for instance, mediate effects that cascade through ecosystems. Trophic interactions in particular, are susceptible to changes in search behaviour - increased search effort increases encounters with prey, but oftentimes also increases risk through increasing encounters with predators (Lima and Dill, 1990), as well as incurring an energetic cost. Evolutionarily consistent search behaviour is thus one that reflects an individual's trade-off between benefit, cost and risk within a particular ecosystem configuration. Further, as the prevailing ecosystem configuration changes, this behaviour may likely adapt so as to maintain this trade-off.

With respect to ecosystem dynamics, a lingering question is how to incorporate behaviour into ecosystem models. That zooplankton change their behaviour is empirically well known, and that changing behaviour has an clear cascading effect on trophic interactions, distribution patterns and population dynamics is also clear. It is almost universally argued, that more empirical observations are needed to prescribe behaviour; behaviour that can then be incorporated into ecosystem model to provide robust predictions on the productivity, structure, biodiversity and sustainability of the marine ecosystem. The underlying premise of this argument though is flawed. Behaviour is not fixed, it is adaptive. While empirical observations may reveal a section of the behavioural responses of a few components of a particular ecosystem, they will never be able to describe the full range over all possible permutations of biotic and abiotic factors. The solution is to realise that behaviour is adaptive for a reason. How organisms perceive their environment, and the behavioural strategies they adopt to perceived opportunities and treats, have been shaped by evolutionary processes to maximise the

fitness of individuals. Thus the behaviour options adopted must also reflect a maximisation of fitness. The link between environmental conditions, behaviour and fitness gives us a toe hold in developing dynamic and evolutionarily consistent models of behaviour and trophic interactions within ecosystems.

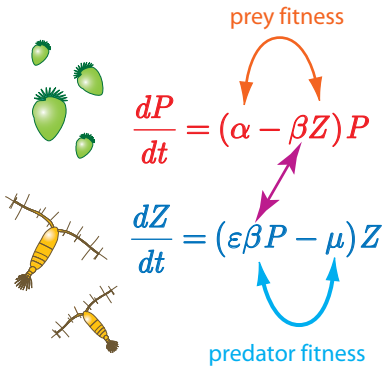


Figure 8.1: The population dynamics of a simple food chain can be described by coupled differential equations – the Lotka Volterra equations. The coupling comes about by a common encounter kernel that describes prey mortality on the one hand, and predator growth on the other. The encounter kernel however in *not* fixed, but is acted on by both predator and prey behaviour as each individual within these populations seeks to optimize its fitness.

A mechanistic understanding from individuals to interactions to populations can be achieved by noting that the same encounter kernel that operates in linking trophic interactions, also appears in evolutionary fitness (at least reproductive success) of the individuals involved. This is illustrated in Fig. 8.1 showing the encounter kernel β of a typical marine predator-prey interaction. These are depicted here as a copepod (Z) and ciliate (P), but may in general be any interaction, *e.g.* bacterium-particle, flagellate-bacterium, larval fish-naplius, big fish-little fish *etc.* The same encounter kernel β enters the population dynamics of both species – regulating the mortality of prey on the one hand and the growth of predators on the other. The encounter kernel also enters into the reproductive success, and therefore evolutionary considerations of both species. Specifically, behaviour is a particularly powerful element in determining the encounter kernel, and it may be expected that adaptive behaviour, sculpted by natural selection, will have an impact on regulating and shaping the structure and functioning of marine ecosystems.

8.1 Prey switching

As a relatively simple example consider a copepod consumer (*e.g. Acartia tonsa*) sandwiched between a predator (*e.g.* larval fish) and a community of protist prey, some of which swim and some of which are immotile. Since some of the copepod's prey are motile, it could adopt an ambush feeding mode, relying on the prey's motility to bring about contacts, thus saving energy costs associated with swimming as well as avoiding the added mortality risk this may have through increasing the probability of contacts with potential predators. Alternatively, the copepod could assume a cruise feeding mode, actively searching for immotile prey, increasing its gross energy intake but also increasing its cost and risk. Many copepods actually exhibit this kind of behaviour, switching between different feeding modes: suspension or cruise feeding on small phytoplankton cells and ambush feeding on motile prey (Tiselius and Jonsson, 1990; Kiørboe et al., 1996; Alcaraz et al., 2007). For a given situation (*e.g.* predator concentration, total prey abundance) the optimal solution, one can suppose, depends on the relative concentration of motile to immotile prey (*cf.* Fig. 8.2). If all the prey are motile, the copepod can afford to be an ambush predator, while when all the prey are immotile, the copepod must swim or face starvation. At some ratio of motile : immotile prey, there is a tipping point – at ratios less than this, fitness is best served by swimming, while at ratios

greater than this, ambush feeding is optimal.

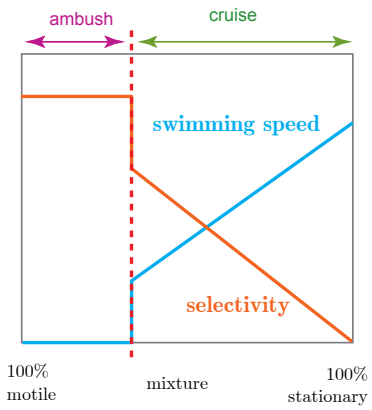


Figure 8.2: A trade-off between benefit, cost and risk predicts that an organisms offered a varying mixture of motile and non-motile prey will switch from ambush to cruise (or suspension) feeding at a critical prey type ratio.

To quantify this, consider first the encounter rate of the copepod with immotile and motile prey respectively:

$$Z_a = \sigma_a C r v \tag{8.1.1}$$

$$Z_b = \sigma_b C (1-r)\sqrt{v^2+u^2} \tag{8.1.2}$$

where C is the total prey concentration irrespective of its motility, r is the proportion of the total prey that is immotile, v is the swimming speed of the copepod, and u is the swimming speed of the motile prey. The net rate of energy intake is thus

$$\dot{E} = \frac{e_a Z_a + e_b Z_b}{1 + h_a Z_a + h_b Z_b} - m - qv^2 \tag{8.1.3}$$

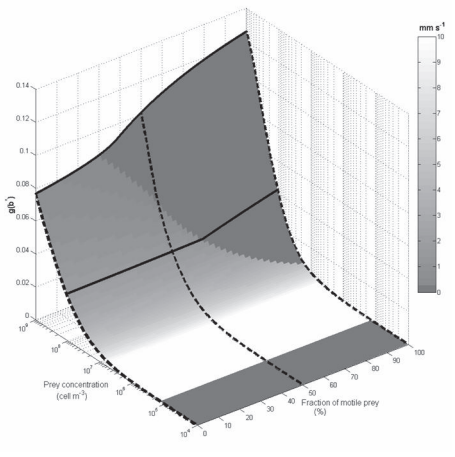


Figure 8.3: Optimal fitness $g(b^*)$ as function of prey concentration and prey type composition expressed as the fraction of motile organisms (0 only static prey; 100 only motile prey with u = 1 mm/s). A shaded map shows the optimal swimming velocity in the range 0 (black) to 10 (white) mm/s has also been superimposed. Solid lines show $g(b*)$ at two prey concentrations C = 10^8 m^{-3} and C = 10^9 m^{-3}. Predator concentration is P = 40 predator/m3. Also shown, the lines (dashed) of 0, 50% and 100% of the fraction of motile prey. Note the switch to zero of the optimal velocity at 65% when C = 10^8 m^{-3} and at 40% when C = 10^9 m^{-3}. Optimal velocity is also zero at low prey concentrations where $g(b^*)$ = 0 (not viable; up to $C \approx 10^5$ m^{-3}.

Larval fish are important predators on copepods. They feed in a pause-travel mode in that they stop periodically to scan a given volume of space for food. If an appropriate food item is found, they attempt to capture and ingest it. After a successful capture, or if no food is

found the larval fish swims on to a new location. An encounter rate model for this type of search may be written as

$$Z_p = \pi R_p^2 C_p v \frac{T_p}{T_s + T_p} + \frac{4}{3}\pi R_p^3 C_p \Phi \frac{1}{T_s + T_p} \tag{8.1.4}$$

Here, T_s an T_p are the mean time spent swimming and paused in a pause-travel cycle. It is assumed in this that the larval fish can only detect a copepod while it is paused. The copepod can be detected either because it was in the predator's scan volume when it paused, or it blunders into the scan volume while the predator is paused. For simplicity I assume a spherical scanning volume of radius R_p but this could be any general shape. The displacement of the fish between pauses may be such that a proportion of scanning volumes overlap. This is expressed by the factor Φ. If larval fish are 100% efficient at capturing detected prey, the net mortality rate for the copepod may thus be written

$$\mu = \mu_0 + Z_p \tag{8.1.5}$$

The optimal fitness $g = \dot{E}/\mu$ is an increasing function of both prey composition as the % motility increases, and as overall prey concentration increases (Fig. 8.3). However, an increase in prey abundance does not always correspond to higher ingestion rate since behavioural switching effects can emerge when predation risk is considered at intermediate and high prey concentrations (Fig. 8.3).

For example at $C = 10^8$ m^{-3}, when the fraction of motile prey falls below 65%, optimal behaviour corresponds to a swimming velocity in the range 2.2 - 1.5 mm/s. As the relative abundance of motile prey items increases, the fitness of the optimal strategy, $g(b^*)$, slowly increases until the optimal velocity switches to zero, *i.e.* copepods switch to an ambush feeding mode with a strong increase in $g(b^*)$ (Fig. 8.3). These results are function of the relative predator and prey abundances.

Therefore, changes in the predator and prey concentrations and/or in the prey composition can induce modification of the optimal swimming behaviour. This in turn produces changes in the realized vital rates of the copepods (*i.e.* their grazing rates, growth and mortality). Fig. 8.4 shows how ingestion (a,b) and mortality (c,d) rates are modified by changes in the swimming speed of the copepod feeding on mixed composition of motile and static prey at two different concentrations. Moving from a cruising into an ambush strategy at lower concentration produces a ca. 40% reduction of the ingestion rates and a ca. 60% reduction of the mortality rate, with a strong increase in the fitness (Fig. 8.4b).

The message is clear: a grazer following an optimal trade-off between foraging and predation risk can significantly influence the transfer rates between trophic levels.

8.2 Simulating ecosystem dynamics

Marine ecosystems are an assemblage of billions of interacting individuals, each attuned to seek an evolutionary advantage. Such a concept is powerful in giving us insights into how and why individuals behave, but it is not immediately useful when it comes determining how an ecosystem will respond to some external factor such as climate change, resource harvesting, or the introduction of an invasive species. Such questions seek to address system level responses (*e.g.* productivity, resilience, stability, diversity) where much of current societal concerns lie. A more manageable picture is to consider marine ecosystems as an assemblage of interacting

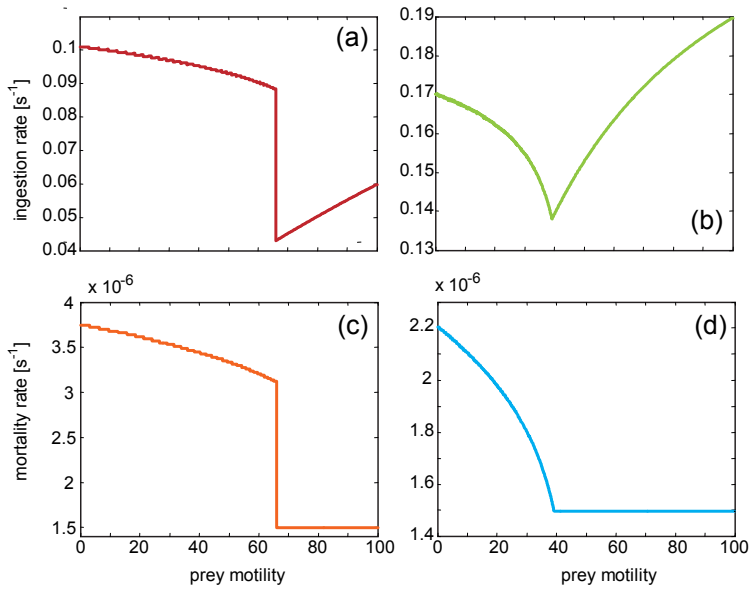


Figure 8.4: Realized ingestion (a,b) and mortality (c,d) rates in *Acartia tonsa* when feeding on mixed prey composition (0 only static prey; 100 only motile prey) at two different concentrations C = 10^8 m^{-3} (a,c) and $C = 10^9$ m^{-3} (b,d). The predator concentration is $P = 40$ m^{-3}.

populations of species. To be sure, there may be thousands of species interlinked in an ecosystem, so the computational and interpretive effort is still formidable. Yet populations yield to rather elegant mathematical renderings that are immediately accessible to analysis.

Even when cast in terms of populations, marine food webs remain complex systems of describe. There are several ways in which this complexity can be reduced and made managable[1]. One such is to organize populations into functional groups depending on the role species have in the flow of energy and material within an ecosystem. One of the simplest rendering of a marine ecosystem is an NPZ model. Phytoplankton (P) fix carbon using sunlight, but are usually limited by nutrients (N). Zooplankton (Z) feed on phytoplankton, and excrete nutrients back into the ecosystem. This gives a rather crude, but effective caricature of an ecosystem, and can be simulated using just a handful of parameters (growth rate, grazing rate, mortality and uptake rate). To make such models more realistic, there are any number of refinements that can be made by adding more functional groups. Other nutrients might be limiting (*e.g.* silicate for diatoms), different sized classes of phytoplankton being grazed on by different sizes of zooplankton, the microbial loop, and the fate of detritus are all elements that have been included over the years. Herein lies a practical problem for with every additional functional group there is a geometric increase in the number of parameters – assuming of course that such parameters can be measured are realistic reflections of community level interactions. Other means of reducing complexity are trait-based models. Body size for instance is a fundamental characteristic of individuals with strong implications on how they interact trophicly with other individuals in an ecosystem: large animals generally eat smaller animals. This, together with some relatively general allometric scaling rules (*e.g.* size depen-

[1] An alternative view, that of embracing complexity, has also been espoused. While it might sound cool, it is an infertile approach as there is no theoretical framework within which it can be achieved. The best that can be hoped for is meta empiricism with little or no predictive capability.

dent base metabolism) produce models from which certain properties of ecosystems emerge – the biomass spectrum for instance. Further, it gives a means of predicting how perturbations on the system will propagate (*e.g.* size selective fisheries mortality) and the ensuing trophic cascades.

Another means is to ignore the full ecosystem configuration and focus in on a particular subset of interactions. This is the concept of community modules, and recognizes that while in general, ecosystems have a baroque topology, often times key species interact in close association with each other. The most basic element of a food web is a tri-trophic community module

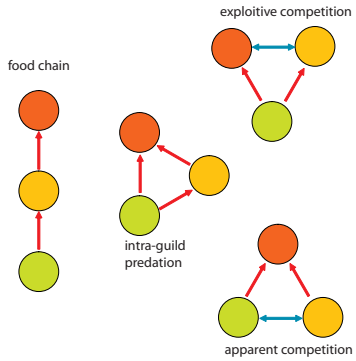


Figure 8.5: Tri-trophic community modules, where 3 species interact either directly or indirectly. In general, there are 4 such arrangements: a linear food chain, 2 predators feeding on a single resource (exploitive competition), a single predator feeding off 2 resources (apparent competition), and intra-guild predation (*e.g.* omnivory) where a top predator can feed on both a resource and a consumer. The red arrows indicate direct (trophic) interactions while the blue arrows indicate indirect interactions

(Fig. 8.5) from which a food-web can be built up. Understanding the dynamics governing such relatively simple arrangements can lend insight into the overall functioning of ecosystems and their emergent properties. It is within this context that I will examine the effect of adaptive behaviour (*i.e.* the foraging – risk trade-off) on populations and communities.

But first some background on population dynamics.

8.2.1 Lotka–Volterra dynamics

Ecosystems are structured through inter-relationships that commonly exhibit feed-back responses. A classical system is that of a predator–prey interaction: prey numbers that would increase exponentially are controlled by a predator which increases in number proportional to the number of prey consumed. When predator numbers become too large however, prey numbers and subsequently predator numbers collapse. The systems exhibit boom and bust cycles. The basic dynamics of these system can be described in a set of coupled differential equations

$$\begin{aligned} \frac{dx}{dt} &= f(x,y) = x\,(\alpha - \beta y) \\ \frac{dy}{dt} &= g(x,y) = y\,(\epsilon\beta y - \mu) \end{aligned} \tag{8.2.1}$$

where x is the prey concentration (number/volume), y is the predator concentration, α is the intrinsic growth rate of the prey (number/time), μ is the intrinsic mortality rate of the predator (number/time), β is the clearance rate of prey by predators (volume/time), and ϵ is a conversion factor – specifically how many prey are required to fuel the generation of a single predator offspring. These are the Lotka-Volterra equations, named after Alfred J. Lotka (1880–1949) and Vito Volterra (1860–1940), who publicized their significance independently in 1925 and 1926 respectively.

There are two steady state solutions to these equations, *i.e.* where $dx/dt = dy/dt = 0$. One is somewhat trivial, in that is is where both $x = 0$ and $y = 0$ – both predator and prey are extinct. The other solution is more interesting and is where

$$\begin{aligned} x &= \frac{\mu}{\epsilon\beta} \\ y &= \frac{\alpha}{\beta} \end{aligned} \tag{8.2.2}$$

These are termed the fixed points of the systems. There is a whole branch of mathematics – dynamical system analysis and its cousins chaos theory and complexity – that examines the behaviour of these types of processes. The fixed points can be visualized as points in predator-prey space (*i.e.* (x, y) space). In general, any dynamic system can be visualized in n-dimensional phase space, where n is the number of variables (*i.e.* populations), and the temporal development of the system can be visualized as a trajectory in this phase space. The underlying behaviour of this system can be determined by examining the Jacobian J in the vicinity of a fixed point. The Jacobian for the above Lotka-Volterra system is given by

$$\mathsf{J} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} \alpha - \beta y & -\beta x \\ \epsilon\beta y & \epsilon\beta x - \mu \end{bmatrix} \tag{8.2.3}$$

The Jacobian says something about how the generator functions $f(x, y)$ and $g(x, y)$ vary locally. The sign of the elements in the Jacobian say something about the direction of system trajectories.

For the first fixed point $(x, y) = (0, 0)$,

$$\mathsf{J}(x, y) = \begin{bmatrix} \alpha & 0 \\ 0 & -\mu \end{bmatrix}$$

which describes a saddle point. What this means is that while $(x, y) = (0, 0)$ is a fixed point, it is unstable. Very close to $(x, y) = (0, 0)$, the prey concentration will increase, taking the system away from the fixed point.

At second fixed point $(x, y) = (\mu/\epsilon\beta, \alpha/\beta)$

$$\mathsf{J}(x, y) = \begin{bmatrix} 0 & -\mu/\epsilon \\ \alpha\epsilon & 0 \end{bmatrix}$$

has eigenvalues $\lambda = \pm i\sqrt{\alpha\mu}$, the real part of which is zero. Thus, the trajectory of the predatory-prey system in phase space in the vicinity of this second fixed-point is a periodic orbit.

8.2.2 Chaos in dynamic systems

In general, any number n of interacting populations may be described as n coupled differential equations that can be concisely written as

$$\frac{d\mathbf{x}}{dt} = \mathsf{A} \cdot \mathbf{x} = \mathsf{F} \tag{8.2.4}$$

where $\mathbf{x}$ is a n-vector describing the number density of each population, and A – termed the community matrix – is a $n \times n$ matrix of interaction terms, that are in general functions of elements of $\mathbf{x}$. The elements of the community matrix are for instance, functions of all the α's, β's, ϵ's and μ's an n dimensional version of the classic 2 dimensional Lotka-Volterra system (Eq. 8.2.1) might contain. The general description of such an inter-connected process is a dynamic system. In general, A and F can be functions of $\mathbf{x}$.

A means of visualizing the dynamics is as the trajectory of the system in phase space – an n dimensional space with the abundance of each population along one axis. The trajectory is solely determined by the dynamics of the system Eq. 8.2.4 and its initial condition $\mathbf{x}_0$. Trajectories can never cross each other in phase space. This is for the simple reason that should the system find itself in exactly the same configuration at two times (*i.e.* at the instant a trajectory supposedly crosses) then by Eq. 8.2.4 its subsequent trajectory will run precisely parallel to the first.

The topology of phase space is organized by the fixed points in the system (*i.e.* where $d\mathbf{x}/dt = 0$), and the behaviour of trajectories in the transient regions between them. Such regions can be "basin-like" attractors, or "hill-like" repellers, or saddles with both attractor and repeller characteristics. They can also belong to limit cycles or quasi-periodic orbits. When the system is non-linear (*i.e.* A is a function of $\mathbf{x}$), and composed of more than two interacting populations (*i.e.* $n > 2$),[2] a region can also be chaotic.

Chaos is an aperiodic long-term behaviour (*i.e.* motion in phase space) arising in a deterministic dynamic system that exhibits a sensitive dependence on initial conditions. Aperiodic long-term behaviour implies that the system does not settle down to a repeating cycle even as time goes to infinity[3]. Deterministic dynamics means there are no random external influences, and all the irregularities in the system arise from its inherent nonlinearities. The sensitivity to initial conditions means that points in phase space (*i.e.* two system configurations that are very nearly identical) will diverge rapidly, irrespective of how close they are initially.

The later condition can be quantified for a system by examining its Lyapunov exponents. These express the local rate at which two points in phase space diverge;

$$\delta\mathbf{x}(t) = e^{\lambda t}\delta\mathbf{x}_0 \tag{8.2.5}$$

For any n dimensional system, there are n Lyapunov exponents for any point in space. They are the eigenvalues of the Jacobian J of the vector F. This can be derived quite simply. A linear expansion in the vicinity of 2 nearby system configurations $\mathbf{x}_a$ and $\mathbf{x}_b$ gives

$$\frac{d\mathbf{x}_a}{dt} = \mathsf{F}(\mathbf{x}_a) = \mathsf{F}(\mathbf{x}_b) + (\mathbf{x}_a - \mathbf{x}_b)\mathsf{J}(\mathbf{x}_a) \tag{8.2.6}$$

But $\mathsf{F}(\mathbf{x}_b)$ is the time rate of change of $\mathbf{x}_b$, so that writing $\delta\mathbf{x} = (\mathbf{x}_a - \mathbf{x}_b)$

$$\frac{d\delta\mathbf{x}}{dt} = \delta\mathbf{x}\mathsf{J} \tag{8.2.7}$$

from which Eq. 8.2.5 follows.

[2] This is true for the continuous time case. In the case of discrete time, chaos can arise out of only one nonlinear relationship.

[3] This is very peculiar especially when one considers that the trajectory of the system cannot cross itself. It is mind-boggling to imagine how these conditions of non-repeating and non-crossing over infinite time can be resolved.

Of particular interest are the Lyapunov exponents in the neighbourhood of the various fixed points of the system. The number of fixed points depends on the dimension of the system, and the level of non-linearity of F. Each fixed point will have n eigenvalues, $\lambda_1, \lambda_2 \ldots \lambda_n$. A fixed point is an attractor if all $Re(\lambda_i) < 0$. However, if a the real part of any of these eigenvalues is positive $Re(\lambda_i) > 0$, the fixed point is unstable (*i.e.* a repellor). If 2 or more of the fixed points are unstable, then the system is chaotic, and a fixed point within a chaotic system with at least one positive, one negative and one zero valued $Re(\lambda_i)$ is a strange attractor.

The behaviour of dynamic systems often depends on the precise value of one or other coefficient in the matrix A. In an ecosystem interpretations, this critical parameter may be the carrying capacity of a base resource, or the trophic efficiency of one of the interaction terms. Increasing a parameter can change the behaviour of a system from stable, to oscillatory to chaotic and back oscillatory, back to chaotic *etc.* This behaviour is revealed in bifurcation diagrams; plots of the value of one variable (x) at a time when some other variable (y) is equal to some fixed value (*i.e.* $y = y_0$). If the system is attracting, then eventually, successive passes will always return the same value of x. If it is a simple oscillation, it will return 2 distinct values of x; or some multiple of 2 for more complex oscillations. If the returned values of x show no coherent pattern, the system is taken to be chaotic – orbits do not repeat themselves. Bifurcation diagrams are a relatively simple illustration of chaos and have thus become popular. A word of warning though. Numerical techniques in integrating dynamic systems are far from simple, and inaccuracies in integration methods may predict chaos where no chaos actually exists.

Instability in dynamic systems need not mean the same as what we take as instability in an ecosystem. Dynamic systems speak of limit cycles, quasi-periodic orbits and chaos as unstable but this does not mean that any of the simulated populations become extinct – as maybe a criterion characterizing the instability of an ecosystem. Simulated population abundances may become vanishingly small, but usually bounce back is some cyclic way. In a practical sense however, real populations cannot become vanishingly small. They are after all composed of individuals that cannot be sub-divided. Further, in the real world, there are random external influences (bad weather, a visit by a migratory predator, failure to find a mate, a sampling expedition by ecologists) the effects of which can be catastrophic when population abundances are low. Systems with large amplitude oscillations whether periodic or chaotic are susceptible to stochastic extinctions, and may be deemed to be ecologically unstable.

8.2.3 Lotka–Volterra competition

An important variant of Lotka-Volterra dynamics is that which considers the competition between closely related species. Suppose these species compete for a single limiting resource, and the environment can only support a finite biomass of all species – *i.e.* the environment has a carrying capacity K. For n inter-competing species, the population dynamics can be written in the n coupled equations

$$\frac{dx_i}{dt} = \alpha_i x_i \left(1 - \frac{\sum_i^n \epsilon_i x_i}{K}\right) - \mu_i x_i \tag{8.2.8}$$

where α_i and μ_i are the intrinsic reproduction rate and mortality rate of species i. The coefficient ϵ_i represents the "occupancy" of an individual of species i of the carrying capacity. While α_i is the intrinsic reproductive rate, the effective rate of reproduction depends also on availability of unoccupied carrying capacity. The system represented here is pathologically

unstable – for any random combinations of $(\alpha_i, \mu_i, \epsilon_i)$, nearly all species will eventually go extinct with only one survivor. The eventual winner of this competition is that species that has the maximal ratio α_i/μ_i. Other species may also survive, but only if they have the same reproductive to mortality rate ratio. In general, one limiting resource will lead to only one winner; n limiting resources will lead to n eventual winners – an observation the has lead to the so-called paradox of the plankton.

While the above is couched in terms of species, it can also be interpreted in terms of individuals – most importantly in terms of "mutant" individuals and their ability to supplant a resident population. Reproduction and mortality rates are generally considered as properties of a population – but they can also be seen as probabilistic properties of individuals. In Eq. (8.2.8) for instance, an individual that through some mutation in ability of behaviour is able to increase its reproduction:mortality ratio over those of its cousins (and thus its immediate competitors) will have a higher probability of increasing its representation in future generations. This is a fairly precise definition of Darwinian fitness. Thus a fairly simple parameter reflecting fitness is

$$g = \frac{\alpha}{\mu} \tag{8.2.9}$$

This is precisely the measure of fitness based on a trade-off between individual growth and mortality introduced in Sec. 7.1.

8.3 Omnivory – an unexpected guild

My particular concern here is how adaptive behaviour affects the dynamics of tri-trophic systems that exhibit a degree of intra-guild predation (Sih and Christensen, 2001). Omnivory falls into this category (Pimm and Lawton, 1978), where a predator feeds off both a consumer and the consumer's resource. Seen purely in the light of direct interactions, such trophic arrangements should be quite rare, as the intra-guild consumer has to endure both direct predation and competition from the top predator. This is borne out in simple dynamic descriptions that suggest that intra-guild predation systems are highly un-stable (Pimm and Lawton, 1978) in their restricted admittance of stable co-existence (Holt, 1996; Holt and Polis, 1997), their prevalence to large amplitude limit cycles (Holt, 1996), and chaotic dynamics (Tanabe and Namba, 2005).

And yet, intra-guild predation is a common feature of all ecosystems.

The effect of adaptive behaviour has been shown to stabilize linear food chains (Abrams, 1992). For tri-trophic systems, adaptive (*i.e.* fitness seeking) behaviour of the omnivore (Krivan and Diehl, 2005), has relatively small effect in facilitating co-existence, whereas the effect of adaptive behaviour of the intra-guild prey can have a more pronounced effect (Okuyama and Ruyle, 2003).

8.3.1 The conceptual arena

Consider a simple planktonic food web, where the resource is a non–motile phytoplankton (*cf.* Fig. 8.6) the carrying capacity of which is set by some limiting nutrient. Both the consumer c and predator p are motile with swimming speed u and v respectively. The trophic interaction between individuals is determined by their encounter rate. For simplicity we consider both the consumer and predator to move ballistically in 3D space. The consumer can detect resource at a distance R_c, while the predator can detect both resource and consumer at

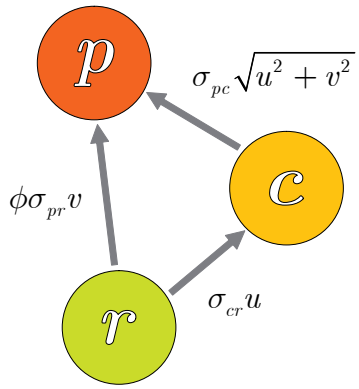


Figure 8.6: A tri-trophic food web with an omnivore top consumer. A resource p (*e.g.* a autotrophic phytoplankton) is fed on by both primary c (*e.g.* ciliate) and a secondary z (*e.g.* copepod) consumers. The secondary consumer is able to feed on both the primary producer and primary consumer. Both c and p are motile with swimming speed u and v respectively.

distance R_p. Further we assume that the resource, consumer and predator are well-mixed (*i.e.* Poisson distributed), that contacts lead to 100% probability of ingestion, and that the handling time is short compared to the mean search time between encounters, *i.e.* we assume a linear functional response. Given these assumptions, the clearance rate of the consumer on the resource is

$$\beta_{cr} = \pi R_c^2 u = \sigma_{cr} u$$

while the clearance rate of predator on resource is

$$\beta_{pr} = \phi\pi R_p^2 v = \phi\sigma_{pr} v$$

Here, ϕ represents the degree of omnivory practiced by the top predator (*i.e.* selectivity of p on r). That is, for $\phi = 0$, the top predator p feeds solely on the consumer c and the trophic interaction becomes a simple linear food chain; for $\phi = 1$, the predator p ingests all resource r encountered. The trophic interaction between predator and consumer depends on the swimming speed of both species. That is

$$\beta_{pr} = \pi R_p^2 \left(u^2 + v^2\right)^{1/2} = \sigma_{pv} w(u)$$

The latter assumes that both the consumer and predator swimming speeds follow a Maxwell distribution, with u and v defined as the mean swimming speed, and w thus being the mean relative speed (*cf.* Sec. 2.1.4).

Provided that all individuals within the resource, consumer and predator populations behave similarly, we can write the population dynamics as a coupled Lotka–Volterra system:

$$\frac{\partial r}{\partial t} = r f_r = r\left\{\rho\left(1 - r/k\right) - \sigma_{cr} u c - \phi\sigma_{pr} v p\right\} \tag{8.3.1}$$

$$\frac{\partial c}{\partial t} = c f_c = c\left\{\varepsilon_{cr}\sigma_{cr} u r - \sigma_{pc} w(u) p - \kappa(u)\right\} \tag{8.3.2}$$

$$\frac{\partial p}{\partial t} = p f_p = p\left\{\varepsilon_{pr}\phi\sigma_{pr} v r + \varepsilon_{pc}\sigma_{pc} w(u) c - \mu_p\right\} \tag{8.3.3}$$

where (r, c, p) are the number density of individuals for resource, consumer and predator species respectively. The carry capacity of the resource is k, σ_{ij} is the search area of species i on species j, while ε_{ij} is the trophic conversion efficiency – the number of individuals of

species j that must be consumed by an individual of species i to produce 1 offspring. The mortality rate of the predator is μ_p, while $\kappa(u)$ is the mortality rate of the consumer – and may be a function of swimming speed u.

This system contains 6 fixed points (*i.e.* where $d(r,c,p)/dt = 0$), one of which denoted $\mathbf{s}_6^* = (r_6^*, c_6^*, p_6^*)$ admit a co-existence of all 3 species. Specifically, all 3 species co-exist provided the 3 conditions hold:

$$r_6^* = k\left\{\frac{\rho\sigma_{pc}\varepsilon_{pc}w + \phi\kappa\sigma_{pr}\varepsilon_{pc}v - \sigma_{cr}u\mu_p}{\rho\sigma_{pc}\varepsilon_{pc}w + k\phi\sigma_{cr}\sigma_{pr}uv\left(\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr}\right)}\right\} > 0 \tag{8.3.4}$$

$$c_6^* = \frac{\mu_p - \phi\sigma_{pr}\varepsilon_{pr}vr_6^*}{\sigma_{pc}\varepsilon_{pc}w} > 0 \tag{8.3.5}$$

$$p_6^* = \frac{\sigma_{cr}\varepsilon_{cr}ur_6^* - \kappa}{\sigma_{pc}w} > 0 \tag{8.3.6}$$

Since all coefficients are $+ve$, the latter 2 conditions can be rewritten as

$$\phi\sigma_{pr}\varepsilon_{pr}vr_6^* < \mu_p \tag{8.3.7}$$

$$\sigma_{cr}\varepsilon_{cr}ur_6^* > \kappa \tag{8.3.8}$$

or

$$\frac{\sigma_{cr}\varepsilon_{cr}u}{\kappa} > \frac{\phi\sigma_{pr}\varepsilon_{pr}v}{\mu_p} \tag{8.3.9}$$

This latter condition can be interpreted as requiring that for a given resource abundance, the consumer must be able to produce more surviving off spring over its expected life time in the absence of predation (κ^{-1}), than the predator can for the same resource abundance over its expected life time (μ_p^{-1}). This is generally stated as the condition that in order for stable co-existence, the consumer c must be able to out-compete the predator p for the resource r.

8.3.2 Optimal swimming speed

We now come to the role of adaptive behaviour. In particular, we assume that the consumer has the ability to adapt its swimming speed in face of a trade-off between encounters with resource (governing its growth rate) and encounters with predators (governing in part its mortality rate). Taking per capita growth rate as a proxy for fitness, the consumer should thus adjust u is such a was so as to maximize the function

$$f_c = \varepsilon_{cr}\sigma_{cr}ru - \sigma_{pc}p\sqrt{u^2+v^2} - \kappa \tag{8.3.10}$$

When the consumer mortality rate is independent of swimming speed ($\kappa(u) = \mu_c$) this has a maximum when the consumer swimming speed is a solution of

$$\frac{u'}{\sqrt{u'^2+v^2}} = \frac{r\varepsilon_{cr}\sigma_{cr}}{p\sigma_{pc}} \tag{8.3.11}$$

However, inserting $u = u'$ into Eq. 8.3.10 returns a negative per capita growth rate (for all allowed values of r and p). That is

$$f_c = \varepsilon_{cr}\sigma_{cr}ru' - \sigma_{pc}p\sqrt{u'^2+v^2} - \mu_c = \frac{-p\sigma_{pc}v^2}{\sqrt{u'^2+v^2}} - \mu_c < 0 \tag{8.3.12}$$

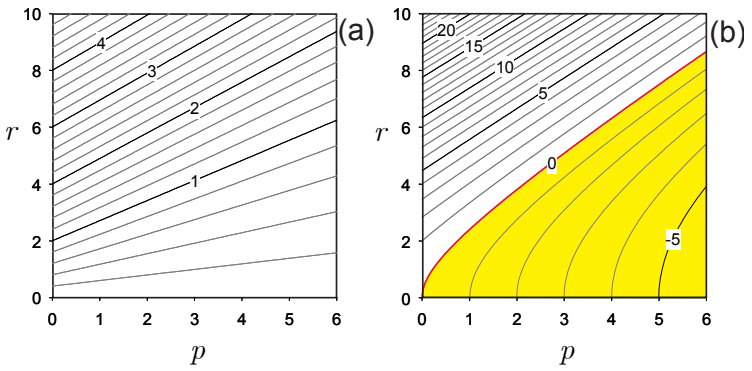


Figure 8.7: Optimal swimming speed and fitness of the consumer c as a function of predator p and prey r concentration within the tri-trophic system.

suggesting co-existence at well defined optimal consumer swimming behaviour (in this rendering) is not possible.

The actual meaning in this model, is that provided

$$r\varepsilon_{cr}\sigma_{cr} > p\sigma_{pc} \tag{8.3.13}$$

the consumer should swim as fast as possible. However, when

$$r\varepsilon_{cr}\sigma_{cr} < p\sigma_{pc} \tag{8.3.14}$$

the consumer should swim at speed u' as given by Eq. 8.3.11 as the least detrimental option. There is a stabilizing feed back in this strategy. In case 8.3.13, swimming as fast as possible increases the grazing rate on r, while also increasing the growth rate of p. Thus r decreases and p increases driving the inequality towards 8.3.14. However, when 8.3.14 applies, c decreases, relieving grazing pressure on r while presenting less of a resource for p, driving the inequality back towards 8.3.13.

To put this adaptive behaviour in a more mechanistic context, we replace the consumer's background mortality rate with a function reflecting the cost of swimming. Specifically we set

$$\kappa(u) = \mu_c + qu^2 \tag{8.3.15}$$

Note that qu^2 can represent either an increase in mortality, or more realistically, a reproductive penalty due to the energetic cost associated with swimming at speed u. Swimming "as fast as possible" now takes on a concrete meaning, and yields

$$2qu' + \frac{p\sigma_{pc}u'}{\sqrt{u'^2 + v^2}} = r\varepsilon_{cr}\sigma_{cr} \tag{8.3.16}$$

where $u'(r,p)$ is the optimal swimming speed of the consumer c for a given resource r and predator p abundance (Fig. 8.7).

The fixed, co-existence point thus becomes

$$r_6^* = k\left\{\frac{\rho\sigma_{pc}\varepsilon_{pc}(u'^2+v^2)^{1/2}+\phi\sigma_{pr}\varepsilon_{pc}vqu'^2-\sigma_{cr}u\mu_p}{\rho\sigma_{pc}\varepsilon_{pc}(u'^2+v^2)^{1/2}+k\phi\sigma_{cr}\sigma_{pr}u'v\left(\varepsilon_{cr}\varepsilon_{pc}-\varepsilon_{pr}\right)}\right\}>0 \tag{8.3.17}$$

$$c_6^* = \frac{\mu_p-\phi\sigma_{pr}\varepsilon_{pr}vr_6^*}{\sigma_{pc}\varepsilon_{pc}(u'^2+v^2)^{1/2}}>0 \tag{8.3.18}$$

$$p_6^* = \frac{\sigma_{cr}\varepsilon_{cr}u'r_6^*-qu'^2}{\sigma_{pc}(u'^2+v^2)^{1/2}}>0 \tag{8.3.19}$$

Noting that u' is a function of the system, the elements of the Jacobian are in general given by

$$J_{ij}=f_i\delta_{ij}+s_j\left\{\frac{\partial f_i}{\partial s_j}+\frac{\partial f_i}{\partial u'}\frac{\partial u'}{\partial s_j}\right\} \tag{8.3.20}$$

where s_j is short hand for r, c or p, and

$$\frac{\partial u'}{\partial r}=\varepsilon_{cr}\sigma_{cr}\left(2q+\frac{p\sigma_{pc}v^2}{(u'^2+v^2)^{3/2}}\right)^{-1} \tag{8.3.21}$$

$$\frac{\partial u'}{\partial p}=-\frac{\sigma_{pc}u'}{(u'^2+v^2)^{1/2}}\left(2q+\frac{p\sigma_{pc}v^2}{(u'^2+v^2)^{3/2}}\right)^{-1} \tag{8.3.22}$$

Given these expressions, it is a relatively straightforward exercise to examine the stability of this system, both for the fixed and adaptive case.

Fig. 8.8 shows the existence and stability of the co-existence fixed point $\mathbf{s}_6$ for both fixed and adaptive dynamics. The free parameters are ϕ, the omnivore selectivity on c, and ε_{pr}, the conversion efficiency of r to p. The parameters are chose to be similar to those in Holt (1996), and Tanabe and Namba (2005) where large amplitude limit cycles and chaotic dynamics are evident. Indeed, for the fixed parameter system, the co-existence fixed point is unstable over much of the parameter space (*cf.* Fig. 8.9). This behaviour turns out to be large amplitude limit cycles with the system trajectory spending much of its time in near the $p=0$ and $r=0$ axes. For the adaptive case however, these oscillations disappear – the co-existence fixed point is stable everywhere within the parameter space.

8.3.3 Stability conditions

What follows is a fairly rigourous proof that the adaptive of the consumer will always lead to a more stable ecosystem. This is somewhat ponderous, as I have yet not found a more elegant approach.

In the spirit of generality, we replace the interaction between c and p with a function $w(u)$, and introduce a general cost-mortality function $\kappa(u)$ on c. That is

$$f_r = \rho\left(1-r/k\right)-\sigma_{cr}uc-\psi\sigma_{pr}vp \tag{8.3.23}$$

$$f_c = \varepsilon_{cr}\sigma_{cr}ur-\sigma_{pc}w(u)p-\kappa(u) \tag{8.3.24}$$

$$f_p = \varepsilon_{pr}\psi\sigma_{pr}vr+\varepsilon_{pc}\sigma_{pc}w(u)c-\mu \tag{8.3.25}$$

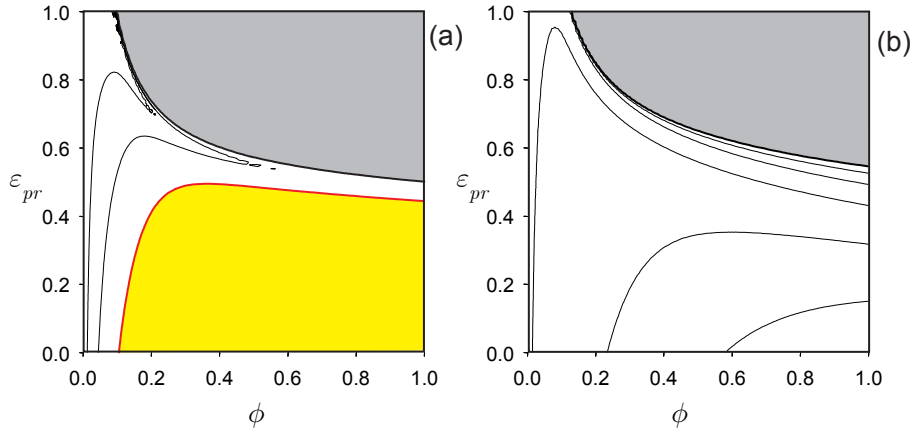


Figure 8.8: The co-existence of $r - c - p$ ($\mathbf{s}_6$) as a function of the degree of omnivory ϕ, and ε_{pr} the trophic conversion efficiency of r to p. Conditions for the co-existence of $\mathbf{s}_6$ is determined. Next, where $\mathbf{s}_6$ exists, the stability of the fixed point is determined from the eigenvalues of the Jacobean. For the fixed case (a) the parameters are $\rho = 1$; $k = 10$; $u = 1$; $v = 1$; $\mu_p = 0.5$; $\varepsilon_{rc} = 1$; $\varepsilon_{cp} = 1$; $\sigma_{rp} = 1$; $\sigma_{rc} = 1$; $\sigma_{cp} = 1/\sqrt{2}$; $q = 0.1$. For the adaptive case (b), the consumer swimming speed u is allowed to vary so as to maximize fitness. Co-existence and stability and are evaluated iteratively. The grey shaded area indicates where c becomes extinct. Co-existence is possible everywhere else. The yellow shaded area in (a) *i.e.* the fixed system, indicates where the fixed point is unstable – describing either limit cycle or chaotic dynamics. For the adaptive case, this unstable region totally disappears – all co-existence points are stable.

where w and κ are increasing functions of u. Optimal swimming speed arises from the condition:

$$\frac{\partial f_c}{\partial u} = \varepsilon_{cr}\sigma_{cr}r - \sigma_{pc}p\frac{\partial w}{\partial u} - \frac{\partial \kappa}{\partial u} = 0 \tag{8.3.26}$$

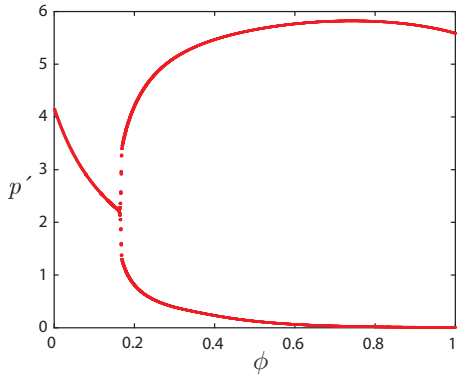


Figure 8.9: The bifurcation diagram for the non-adaptive system described above, for wich the stability condition is plotted in Fig. 8.8.b. The diagram plots the maximum and minimum values of p for various values of ϕ and with $\varepsilon_{pr} = 0.2$. The dynamics change from a stable, stationary solution to a simple period orbit at $\phi = 0.19$. Minimum values of p as $\phi \to 1$ remain finite but become exceedingly small giving rise to potential stochastic extinction.

so that in general $u(r,p)$. By differentiating Eq. 8.3.26 with respect to r and p we get

$$\frac{\partial u}{\partial r} = \varepsilon_{cr}\sigma_{cr}\left\{p\frac{\partial^2 w}{\partial u^2}+\frac{\partial^2 \kappa}{\partial u^2}\right\}^{-1} \tag{8.3.27}$$

$$\frac{\partial u}{\partial p} = -\sigma_{pc}\frac{\partial w}{\partial u}\left\{p\frac{\partial^2 w}{\partial u^2}+\frac{\partial^2 \kappa}{\partial u^2}\right\}^{-1} \tag{8.3.28}$$

which combined give

$$\varepsilon_{cr}\sigma_{cr}\frac{\partial u}{\partial p}+\sigma_{pc}\frac{\partial w}{\partial u}\frac{\partial u}{\partial r}=0 \tag{8.3.29}$$

Thus we have from Eq. 8.3.27 and Eq. 8.3.28 that provided:

$$\left\{\frac{\partial w}{\partial u}>0,\frac{\partial^2 w}{\partial u^2}\geq 0,\frac{\partial^2 \kappa}{\partial u^2}\geq 0\right\} \tag{8.3.30}$$

the optimal swimming speed u is an increasing function of r, and a decreasing function of p. Eq. 8.3.30 states our first set of assumption on the general characteristics of the system.

The stability of the adaptive system can be examined using $\mathbf{J}'$, the Jacobian for the adaptive model evaluated at the co-existence fixed point $\mathbf{s}_6'$. Taking into account the functional dependence $u(r,p)$, this can be written

$$\mathbf{J}'=\begin{pmatrix} r\left(\frac{\partial f_r}{\partial r}+\frac{\partial f_r}{\partial u}\frac{\partial u}{\partial r}\right) & r\frac{\partial f_r}{\partial c} & r\left(\frac{\partial f_r}{\partial p}+\frac{\partial f_r}{\partial u}\frac{\partial u}{\partial p}\right) \\ c\frac{\partial f_c}{\partial r} & 0 & c\frac{\partial f_c}{\partial p} \\ p\left(\frac{\partial f_p}{\partial r}+\frac{\partial f_p}{\partial u}\frac{\partial u}{\partial r}\right) & p\frac{\partial f_p}{\partial c} & p\frac{\partial f_p}{\partial u}\frac{\partial u}{\partial p} \end{pmatrix} \tag{8.3.31}$$

while that for the non-adaptive model, $\mathbf{J}$, has an identical form, somewhat simplified with $\partial u/\partial r = 0$ and $\partial u/\partial p = 0$. For these systems, the co-existence fixed point is an attractor provided their respective eigenvalues all have negative real parts. Indeed, from the signs of the factors involved, we can immediately say that

$$tr(\mathbf{J}')<tr(\mathbf{J})<0 \tag{8.3.32}$$

Since the sum of eigenvalues equals the trace of the matrix, those of the adaptive system will tend to be more negative (specifically negative over a greater range of parameters) than those for the non-adaptive system. This is a result that is at least consistent with, although it in no way proves, the stabilizing effects of adaptive behaviour on system dynamics.

We can determine the general stability of the system by firstly examining the characteristic polynomial of $\mathbf{J}$ for non-adaptive behaviour,

$$\lambda^3+\alpha\lambda^2+\beta\lambda+\gamma=0 \tag{8.3.33}$$

where

$$\alpha = -r\frac{\partial f_r}{\partial r} \tag{8.3.34}$$

$$\beta = -cp\frac{\partial f_c}{\partial p}\frac{\partial f_p}{\partial c}-rc\frac{\partial f_r}{\partial c}\frac{\partial f_c}{\partial r}-rp\frac{\partial f_r}{\partial p}\frac{\partial f_p}{\partial r} \tag{8.3.35}$$

$$\gamma = -rcp\left\{\frac{\partial f_r}{\partial c}\frac{\partial f_c}{\partial p}\frac{\partial f_p}{\partial r}+\frac{\partial f_r}{\partial p}\frac{\partial f_p}{\partial c}\frac{\partial f_c}{\partial r}-\frac{\partial f_r}{\partial r}\frac{\partial f_c}{\partial p}\frac{\partial f_p}{\partial c}\right\} \tag{8.3.36}$$

The fixed point $\mathbf{s}_6$ is stable when the 3 roots of the characteristic polynomial all have negative real parts, a condition that is met provided $\alpha, \beta, \gamma > 0$ and $\alpha\beta > 0$. This is a consequence of the Routh-Hurwitz stability criterion. In general, the coefficients α and β are always positive. Furthermore, because the f functions display certain symmetries:

$$\frac{\partial f_c}{\partial r} = -\varepsilon_{cr}\frac{\partial f_r}{\partial c}; \quad \frac{\partial f_p}{\partial c} = -\varepsilon_{pc}\frac{\partial f_c}{\partial p}; \quad \frac{\partial f_p}{\partial r} = -\varepsilon_{pr}\frac{\partial f_r}{\partial p} \tag{8.3.37}$$

γ becomes

$$\gamma = -rcp\frac{\partial f_c}{\partial p}\left\{\frac{\partial f_r}{\partial c}\frac{\partial f_r}{\partial p}(\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr}) + \frac{\partial f_r}{\partial r}\frac{\partial f_c}{\partial p}\varepsilon_{pc}\right\} \tag{8.3.38}$$

This can be confirmed by checking the signs of the various derivatives involved.

Thus γ is always positive provided

$$\varepsilon_{cr}\varepsilon_{pc} > \varepsilon_{pr} \tag{8.3.39}$$

This is our second general assumption.

The controlling condition for stability of the non-adaptive thus reduces to $\alpha\beta > \gamma$.

We now turn our attention to the stability criterion for the adaptive system, where the Jacobian $\mathbf{J}'$ has the characteristic polynomial

$$\lambda^3 + \alpha'\lambda^2 + \beta'\lambda + \gamma' = 0 \tag{8.3.40}$$

It is convenient to examine the differences between the adaptive and non-adaptive system. These are given by

$$\alpha' - \alpha = -r\frac{\partial f_r}{\partial u}\frac{\partial u}{\partial r} - p\frac{\partial f_p}{\partial u}\frac{\partial u}{\partial p} \tag{8.3.41}$$

$$\beta' - \beta = rp\left\{\frac{\partial f_r}{\partial r}\frac{\partial f_p}{\partial u}\frac{\partial u}{\partial p} - \frac{\partial f_p}{\partial r}\frac{\partial f_r}{\partial u}\frac{\partial u}{\partial p} - \frac{\partial f_r}{\partial p}\frac{\partial f_p}{\partial u}\frac{\partial u}{\partial r}\right\} \tag{8.3.42}$$

$$\gamma' - \gamma = rcp\left\{\frac{\partial f_p}{\partial u}\frac{\partial f_r}{\partial c} - \frac{\partial f_r}{\partial u}\frac{\partial f_p}{\partial c}\right\}\left\{\frac{\partial f_c}{\partial r}\frac{\partial u}{\partial p} - \frac{\partial f_c}{\partial p}\frac{\partial u}{\partial r}\right\} \tag{8.3.43}$$

From the signs of gradients involved, we can immediately write $\alpha' - \alpha > 0$.

We can also write the following useful identities:

$$\frac{\partial f_r}{\partial u} = -\sigma_{cr}c \tag{8.3.44}$$

$$\frac{\partial f_p}{\partial u} = \varepsilon_{pc}\sigma_{pc}c\frac{\partial w}{\partial u} \tag{8.3.45}$$

These together with Eq. 8.3.46 give

$$\frac{\partial u}{\partial r}\frac{\partial f_p}{\partial u} = \varepsilon_{cr}\varepsilon_{pc}\frac{\partial u}{\partial p}\frac{\partial f_r}{\partial u} \tag{8.3.46}$$

Substituting into Eq. 8.3.42 gives

$$\beta' - \beta = rp\frac{\partial u}{\partial p}\left\{\frac{\partial f_r}{\partial r}\frac{\partial f_p}{\partial u} - \frac{\partial f_r}{\partial p}\frac{\partial f_r}{\partial u}(\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr})\right\} \tag{8.3.47}$$

and is always positive provided the condition already encountered in Eq. 8.3.39 holds.

Likewise, it can be readily shown

$$\frac{\partial u}{\partial p}\frac{\partial f_c}{\partial r} = \frac{u}{w}\frac{\partial w}{\partial u}\frac{\partial u}{\partial r}\frac{\partial f_c}{\partial p} \tag{8.3.48}$$

$$\frac{\partial f_p}{\partial u}\frac{\partial f_r}{\partial c} = \frac{u}{w}\frac{\partial w}{\partial u}\frac{\partial f_r}{\partial u}\frac{\partial f_p}{\partial c} \tag{8.3.49}$$

which when substituted into Eq. 8.3.43 gives

$$\gamma' - \gamma = rcp\frac{\partial u}{\partial r}\frac{\partial f_r}{\partial u}\frac{\partial f_c}{\partial p}\frac{\partial f_p}{\partial c}\left\{\frac{u}{w}\frac{\partial w}{\partial u} - 1\right\}^2 \tag{8.3.50}$$

Thus, $\gamma' - \gamma$ is also always positive.

Summing up we have

$$\begin{aligned} 0 &< \alpha &< \alpha' \\ 0 &< \beta &< \beta' \\ 0 &< \gamma &< \gamma' \end{aligned} \tag{8.3.51}$$

Thus the stability of either system is determined by the respective conditions $\alpha\beta > \gamma$ and $\alpha'\beta' > \gamma'$. Whatever the configuration of the non-adaptive system, we can assert that adaptation enhances stability provided $\alpha'\beta' - \gamma' > \alpha\beta - \gamma$.

Expanding out this expression, we get

$$\begin{aligned} & \alpha'\beta' - \alpha\beta - (\gamma - \gamma') \\ = \; & c\left(\overset{(-)}{\frac{\partial f_c}{\partial p}}\overset{(+)}{\frac{\partial f_p}{\partial u}}p + \overset{(+)}{\frac{\partial f_c}{\partial r}}\overset{(-)}{\frac{\partial f_r}{\partial u}}r\right)\left(\overset{(+)}{\frac{\partial f_p}{\partial c}}\overset{(-)}{\frac{\partial u}{\partial p}}p + \overset{(-)}{\frac{\partial f_r}{\partial c}}\overset{(+)}{\frac{\partial u}{\partial r}}r\right) \\ + \; & rp\overset{(-)}{\frac{\partial f_r}{\partial p}}\left(\overset{(+)}{\frac{\partial f_p}{\partial r}} + \overset{(-)}{\frac{\partial f_r}{\partial u}}\overset{(-)}{\frac{\partial u}{\partial p}}(\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr})\right)\left(\overset{(+)}{\frac{\partial f_p}{\partial c}}\overset{(-)}{\frac{\partial u}{\partial p}}p + \overset{(-)}{\frac{\partial f_r}{\partial c}}\overset{(+)}{\frac{\partial u}{\partial r}}r\right) \\ + \; & rp\overset{(-)}{\frac{\partial f_r}{\partial r}}\overset{(-)}{\frac{\partial u}{\partial p}}\left(r\overset{(-)}{\frac{\partial f_r}{\partial p}}\overset{(-)}{\frac{\partial f_r}{\partial u}}(\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr}) - \overset{(+)}{\frac{\partial f_p}{\partial u}}\left(\overset{(+)}{\frac{\partial f_p}{\partial c}}\overset{(-)}{\frac{\partial u}{\partial p}}p + \overset{(-)}{\frac{\partial f_r}{\partial c}}\overset{(+)}{\frac{\partial u}{\partial r}}r\right)\right) \\ - \; & r^2p\left(\overset{(-)}{\frac{\partial f_r}{\partial r}}\right)^2\overset{(+)}{\frac{\partial f_p}{\partial u}}\overset{(-)}{\frac{\partial u}{\partial p}} + \end{aligned}$$

where I have written the sign of the various contributions as an aid. Thus, provided $\varepsilon_{cr}\varepsilon_{pc} - \varepsilon_{pr} > 0$, a condition already encountered, $\alpha'\beta' - \alpha\beta - (\gamma - \gamma') > 0$, which together with the proof of Eq. 8.3.51, establishes the stabilizing influence of the adaptive behaviour on tri-trophic systems.

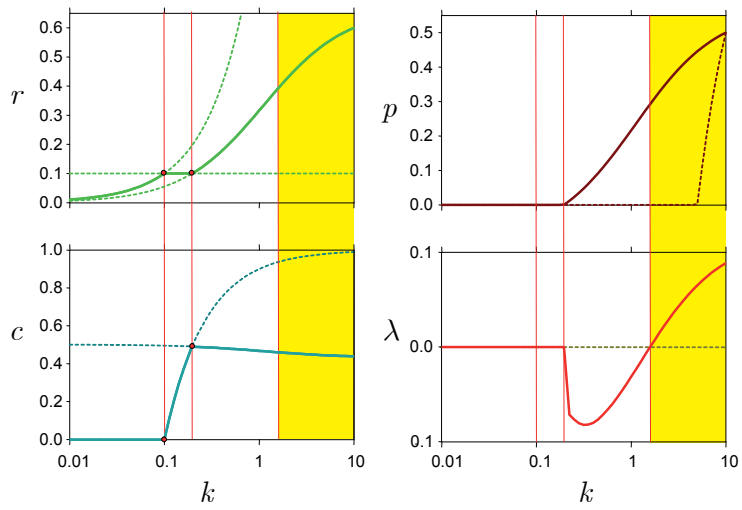


Figure 8.10: The enrichment problem for the fixed tri-trophic system. As the carrying capacity k increases, first r increases with c and p excluded. Beyond a threshold k_c, c appears, followed by p at some higher threshold k_p. As k increases even further however, λ the real part of the largest eigenvalue becomes positive. The region shaded yellow indicate an unstable community configuration.

8.3.4 Enrichment

Simple ecological models along the lines of multi-species Lotka-Volterra interactions predict that enrichment – the increase of primary resources or carrying capacity – can reduce ecosystem stability by promoting large amplitude limit cycles and chaos. This is both counter intuitive (that promoting growth somehow can lead to extinction), and counter to much empirical evidence, and has lead some to term this the paradox of enrichment. (Rosenzweig1971).

If, for instance, we take the same system as described in Eq. 8.3.1, 8.3.1 and 8.3.3, (*i.e.* non-adaptive dynamics)and slowly increase the carrying capacity k, then the resource species can increases linearly $r = k$, but c and p are excluded. Increasing still further to a threshold

$$k_c = \frac{\kappa}{\varepsilon_{cr}\sigma_{cr}u} \tag{8.3.52}$$

the consumer c becomes established. As k increases further, the predator p can invade. This occurs at the threshold

$$k_p = \frac{k_c}{1 + \dfrac{\sigma_{cr}u(\mu - \phi\sigma_{pr}v)}{w\rho\varepsilon_{pc}}} \tag{8.3.53}$$

This is illustrated in Fig. 8.10. As the carrying capacity increases even further, both r and p increase while c decreases slightly. Eventually, the system becomes unstable as witnessed by the positive real part of the largest eigenvalue.

In comparison, when the consumer c is allowed to adapt its behaviour in response to a foraging fitness trade-off, this instability disappears. This is illustrated in Fig. 8.11.

The paradox of enrichment can thus be resolved through adaptive foraging behaviour.

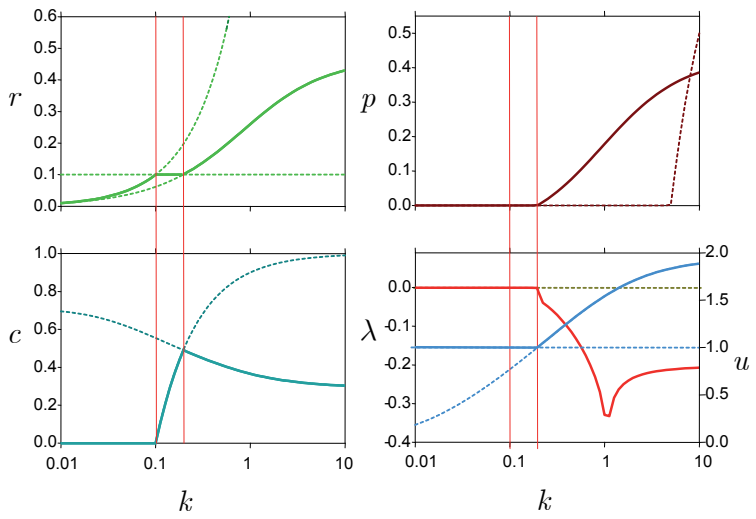


Figure 8.11: The enrichment problem for the adaptive tri-trophic system. As the carrying capacity k increases, first r then c and then p appear, much as for the fixed tri-trophic system. However, the community configuration remains stable (*i.e.* attracted to a fixed point) even at extremely high values of the carrying capacity. The the real part of the largest eigenvalue λ (in red) remains negative. The optimal swimming speed u (in blue) slowly increases

Adaptive behaviour plays a potentially important role in stabilizing ecosystems. In a simple sense, this can be deduced from the instantaneous (or at least rapid) response of behavioural adjustments to opportunity and hazard (trade-offs between benefit-cost-risk) compared to numerical responses (birth and death) to potential boom and bust cycles in resources and predators. Behaviourally-mediated interaction strengths between populations vary in the same sense as numerical interactions, but without the reproductive time lag, so effectively damping oscillations and promoting stability.

Adaptive behaviour not only has implications for individuals, but also for population dynamics. Changing behaviour has cascading effects across trophic levels; impacting not only the growth rate and mortality of an organism, but also the mortality of its prey and the growth rate of its predators.

In particular, an organism adapting its foraging behaviour to its own evolutionary self-interest, can stabilize a community that would otherwise exhibit unstable dynamics (*i.e.* large amplitude oscillations and chaos).

Omnivory is an example of a community module that, for non-adaptive dynamics, should be rare in nature. Its observed prevalence is thus somewhat of a paradox. This paradox can be resolved by appealing to adaptive behaviour which leads to much more stable community dynamics.

9 Concluding Remarks

Marine ecosystems are complex. They are composed of innumerable individual members of many different species interacting with each other through multiple relationships. At its very base, marine ecosystems are driven by the synthesis of organic carbon, most commonly by phytoplankton using sunlight, and to a lesser degree by certain strains of bacteria utilizing the energy inherent in chemical potentials. Organic carbon subsequently fuels all manner of secondary consumers. These are the zooplankton; from the microbes that feed on leaking organic compounds to protist and copepod grazers and the vast swarms of krill that feed the Southern Ocean. They include drifting tentacled forms, gelatinous house builders, as well as the early life stages of molluscs, crustaceans and fishes. Zooplankton themselves are the targets of predation, mostly as it turns out from amongst their own members, but also by fish and mammals. The baroque pattern of who-eats-who that result is far from linear; its has loops and side-way shunts where small-eats-big, where microbial colonizer of detrital accretions found mini-ecosystems, where parasitism and omnivory are common feeding guilds, and where mutual benefits can be afforded through distant trophic relationships. Neither is this pattern constant in time. The daily light cycle for instance imprints a rhythm through visual predators, and the seasonal changes in primary productivity in high latitude and temperate seas sends shock waves through marine ecosystems. More subtle still are the changing feeding behaviours many zooplankton exhibit, as certain prey types are abandoned in favour of others that may be more numerous, more nutritious, more easily caught or less risky. That is, feeding behaviours change as individual members of the ecosystem seek to find the best solution to meeting life's requirements given the conditions they find themselves in.

Thus, like Darwin's tangled bank, marine ecosystems are complex. Indeed complexity is a general hallmark of biological systems; the intricate bio-chemistry of DNA, the design of an individual organism, and the structure of plankton communities, all exhibit an interplay between multiple and different components that somehow conspire to produce a system that is "greater than the sum of its parts". How to approach complexity? Somewhere in here, the philosophical debate on holism versus reductionism always enters in (holism good, reductionism bad), but its nuances are lost on me, and I would suggest that it should be ignored by anyone seeking an actual solution. The phrase "greater than the sum of its parts" can be easily misconstrued to imply that it is more than just the component parts that are active in determining the system. But there is nothing else, just the components and how the relate to each other, and if not open to deep-down reductionist dissection to molecular, quanta, and quarks, they are certainly amendable to a modular version (*i.e.* hierarchical reductionism as espoused by Richard Dawkins). A complex system may be greater than the *sum* of its parts, but can not be greater than the the *product* of its parts.

Within the study of ecosystem complexity, a concept that is well worth exploring is that complex patterns emerge from simple rules. This is perhaps not universally true, but it

certainly is appealing if for no other reason than Occam's razor. For instance, the spatial patterns evident in Lévy walks can emerge from much simpler, and ecologically relevant rules of searching for patchy resources and the dodge-and-weave path geometry of foraging organisms emerges from a trade off between search efficiency and exposure to predation risk. More importantly, there is only a handful of general rules that dictate the behaviour of marine organisms that can be summarized as: (1) find food, (2) reproduce and (3) avoid getting eaten. These may be expressed in any number of ways, but their generality can be surmised from their evolutionary relevance. Indeed, a behaviour that in some way successfully balances all 3 tasks in the face of environmental pressures will give the organism an evolutionary advantage, *i.e.* a high fitness rating.

Information is the handmaiden of behaviour. That is, an appropriate behaviour requires some input of information regarding the state of the individual organism, the environment it is subject to, and a set of rules (algorithms if you will) as to what to do. Indeed zooplankton have invested heavily in sensory equipment; shear probes and stress sensors to detect fluid motion, chemical sensors to detect molecules from the mundane taste of their prey to the exotic allure of pheromones, and while only a few are equipped with imaging eyes, many have a sensory spot that gives rudimentary information on light. Information is a vital resource on a par with energy and mating opportunities. The other crucial ingredient is some means of transducing information into action; given all the possible responses to a given situation, some will promote the organism's reproductive success while others won't. This information is encoded in the organism's blue print – its DNA – and reflects the successful choices of legions of ancestors confronted by similar situations in the past. Behavioural algorithms are phenotypic characteristics in exactly the same way as physical adaptations such as cunning camouflage, sharp predatory claws or strong flight muscles. Indeed, many physical adaptations would have limited use unless accompanied by behavioural adaptations: the efficient musculature of copepods used in their spectacular escape jumps become effective when accompanied with both sensory apparatus and information processing abilities that can translate perceived signals into appropriate action. Precisely how behavioural algorithms work is a trade-off in complexity in itself. Chemical "carrots and sticks" triggered by both internal and external signals goad the organism to action. You just have to look to our own reliance on "gut-feelings" in making decisions to see how this works. Be that as it may, in a modular form, we can consider behaviour algorithms enacted by organisms as traits (think of little segments of computer code), undergoing the same design process (re-programming) through natural selection as their physical form.

Thus, given a means of assessing evolutionary imperatives and information acquisition, we can peer into the rationale of why organisms do what they do, not just qualitatively but qualitatively as well. The quantification of behaviour is of paramount importance not simply in that it allows for hypothesis testing and prediction of individual behaviour, but that it also provides a means of mechanistically deriving the interaction rates between plankton populations. That is, the interaction rates between populations emerge from the behaviour of individuals members of these populations which in turn emerges from the benefit-cost-risk trade-offs each individual member is confronted with. Importantly, these trade-offs change from time to time, so that realized behaviour adapts to circumstance and population interactions are dynamic. It is this adaptive behaviour that has a powerful effect on the structuring and functioning of marine ecosystems. It elicits behaviourally mediated trophic cascades, cements omivory as a stabilizing influence and smooths out large amplitude oscillations in ecosystems coming under resource enrichment. It is, in part, the cunning strategy Robert

May alludes to, through which life conspires to develop complex, highly diverse and stable ecosystems.

In this work I have promoted a very simple fitness measure that tries to capture the major trade-off that all organisms seek to optimize, which may be summarized as

$$\text{fitness} = \frac{\text{benefit} - \text{cost}}{\text{risk}} \tag{9.0.1}$$

Here benefit and cost can be expressed in terms of energy, and risk in terms of mortality. Its appeal, at least to my mind, rests in its simplicity, and yet it can give rise to a rich tapestry of trade-offs (complex patterns from simple rules). Further it is mechanistic. Benefit, cost and risk can each be expressed and quantified in terms of one or other behavioural option; swimming speed, path geometry, time allocation, prey switching and vertical migration being the few examples I have explored in this work. Many others such as over wintering strategies, growth-reproduction allocation, and reproductive behaviour can be explored by similar means. I am willing to concede that the expression 9.0.1 is in all likelihood not a complete and comprehensive measure, and in many instances it may well be wrong. But in this, I hope that I am wrong in a useful way giving impetus to better, and more robust measures.

Finally, the concepts I have developed here have a direct implication for the way we view the future of the marine environment. As the oceans come under increasing pressure from climate change and resource harvesting, we need more reliable models that can predict how the oceans and the life they support will respond. We can no longer rely on our empirical knowledge of the recent past as the entire earth–ocean–climate–biosphere enters a new configuration and the shortcomings of physico-chemical view of marine ecosystems become ever more apparent. By mobilizing concepts from evolutionary ecology, we have extra information from which a mechanistic description of the interaction between marine organisms can be derived. An important aspect of this is that fundamental trophic interactions and the vital rates that depend on them (*viz.* growth, reproduction and mortality) can be viewed as dynamic properties that emerge from the evolutionary imperatives of living organisms. A new generation of marine ecosystem models presents itself, where biodiversity rather than being imposed by model architecture, is an emergent property, and through which new insights are shed into the pressing questions that society asks of marine scientists and oceanographers concerning the substantiality, resilience and species-richness of the future oceans.

Bibliography

Abraham ER. The generation of plankton patchiness by turbulent stirring. *Nature*, 391:577–580, 1998.

Abrams PA. Predators that benefit prey and prey that harm predators: Unusual effects of interacting foraging adaptations. *Am Nat*, 140(4):573–600, 1992.

Alcaraz M, Saiz E, Marras C, and Vaqu D. Effects of turbulence on the developmenrt of phytoplankton biomass and copepod populations in marine microcosms. *Mar Ecol Prog Ser*, 49:117–125, 1988.

Alcaraz M, Saiz E, and Calbet A. Centropages behaviour: Swimming and vertical migration. *Prog Oceanog*, 72:121–136, 2007.

Auton T, Hunt J, and Prudhomme M. The force exerted on a body in inviscid unsteady nonuniform rotational flow. *J. Fluid Mech.*, 9:241–257, 1988.

Babiano A, Cartwright JHE, Piro O, and Provenzale A. Dynamics of a small neutrally buoyant sphere in a fluid and targeting hamiltonian systems. *Phys Rev Lett*, 84(25):5764–5767, 2000.

Bagøien E and Kiørboe T. Blind dating - mate finding in planktonic copepods. i. tracking the phermone trail of *Centropages typicus*. *Mar Ecol Prog Ser*, 300:105–115, 2005a.

Bagøien E and Kiørboe T. Blind dating - mate finding in planktonic copepods. iii. hydromechanical communication in *Acartia tonsa*. *Mar Ecol Prog Ser*, 300:129–133, 2005b.

Bartumeus F. Lévy processes in animal movement: An evolutionary hypothesis. *Fractals-Complex Geometry Patterns and Scaling in Nature and Society*, 15(2):151–162, 2007.

Bartumeus F, Peters F, Pueyo S, C. M, and Catalan J. Helical lévy walks: Adjusting searching statistics to resource availability in microzooplankton. *Proc Nat Acad Sci USA*, 100(22):12771–12775, 2003.

Basset AB. On the motion of a sphere in a viscous liquid. *Philos. Trans. R. Soc. London Ser. A*, 179: 43–63, 1888.

Batchelor GK. *An Introduction to Fluid Dynamics*. Cambridge University Press, Cambridge, 1967.

Batchelor G. The effect of homogeneous turbulence on material lines and surfaces. *Proc R Soc Lond A*, 213(1114):349–366, 1952.

Batchelor G. Small scale variation of convected quantities like temperature in a turbulent fluid. part 1. general discussion and the case of small conductivity. *J Fluid Mech*, 6:113–133, 1959.

Becker LE, Koehler SA, and Stone HA. On self-ptopulsion of micro-machines at low reynolds number: Purcell's three-hinge swimmer. *J Fluid Mech*, 490:15–35, 2003.

Behrenfeld MJ and Falkowski PG. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol Oceanogr*, 42:1–20, 1997.

Benhamou S. How many animals realy do the lévy walk? *Ecology*, 88(8):1962–1969, 2007.

Berg HC. *Random walks in biology*, volume 2. Princeton University Press, Princeton, 1992.

Berg HC and Brown DA. Chemotaxis in escherichia coli analysed by three-dimensional tracking. *Nature*, 239:500–504, 1977.

Blake JR. A spherical envelope approach to ciliary propulsion. *J Fluid Mech*, 46:199–208, 1971.

Blake JR. A model for the micro-structure in cilated organisms. *J Fluid Mech*, 55:1–23, 1972.

Boussinesq VJ. Sur la resistance qu'oppose un liquide indéfini en repos. *C. R., Acad. Sci.,*, 100: 935–937, 1885.

Bundy MH, Gross TF, Vanderploeg HA, and Strickler JR. Perception of inert particles by calanoid copepods: behavioral observations and a numerical model. *J Plankton Res*, 20:2129–2152, 1998.

Buskey EJ. Swimming pattern as an indicator of the roles of copepod sensory systems in the recog-

nition of food. *Mar Biol*, 79:165–175, 1984.
Buskey EJ. Components of mating behavior in planktonic copepods. *J Mar Syst*, 15(1-4):13–21, 1998.
Buskey EJ and Stoecker DK. Locomotory patterns of the planktonic ciliate *Favella* sp.: Adaptations for remaining within food patches. *Bull Mar Sci*, 43(3):783–796, 1988.
Buskey EJ, Coulter C, and Strom S. Locomotory patterns of microzooplankton: Potential effects on food selectivity of larval fish. *Bull Mar Sci*, 53(1):29–43, 1993.
Caparroy P, Perez M, and Carlotti F. Feeding behaviour of *Centropages typicus* in calm and turbulent conditions. *Mar Ecol Prog Ser*, 168:109–118, 1998.
Caparroy P, Thygesen UH, and Visser AW. Modelling the attack success of planktonic predators: patterns and mechanisms of prey size selectivity. *J Plankton Res*, 22(10):1871–1900, 2000.
Cassie RM. Microdistribution of plankton. *Oceanogr Mar Biol Ann Rev*, 1:223–252, 1963.
Chandrasekhar S. Stochastic problems in physics and astronomy. *Rev Mod Phys*, 15:1–89, 1943.
Chapman S. On the brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid. *Proc R Soc Lond A*, 119:34–54, 1928.
Charlesworth B. *Evolution in Age-structured Populations*. Cambridge University Press, Cambridge, UK, 1994.
Charnov EL. Optimal foraging: the marginal value theorem. *Theor Popul Biol*, 30:45–75, 1976.
Childress S. *Mechanics of swimming and flying*. Cambridge University Press, Cambridge, 1981.
Childress S, Koehl MAR, and Miksis M. Scanning currents in stokes flow and the efficient feeding of small organisms. *J Fluid Mech*, 177:407–426, 1987.
Chwang AT and Wu TY. A note on the helical movement of micro-organisms. *Proc R Soc Lond A*, 178:327–346, 1971.
Clift R, Grace JR, and Webber ME. *Bubbles, Drops and Particles*. Academic Press, San Diego, 1978.
Cocke W. Turbulent hydrodynamic line stretching: consequences of isotropy. *Phys Fluids*, 12:2488–2492, 1969.
Costello JH, Strickler JR, Marras C, Trager G, Zeller R, and Freise AJ. Grazing in a turbulent environment: behavioural response of the calanoid copepod centropages hamatus. *Proc Natl Acad Sci USA*, 87:1648–1652., 1990.
Crenshaw HC. A new look at locomotion in microorganisms: rotating and translating. *Am Zool*, 36: 608618, 1996.
Darwin C. Note on hydrodynamics. *Proc. Camb. Phil. Soc.*, 49:342354, 1953.
Davis CS, Flierl GR, Wiebe PH, and Franks PJS. Micropatchiness, turbulence and recruitment in plankton. *J Mar Res*, 49:109–151, 1991.
DeMott WR and Watson MD. Remote detection of algae by copepods: response to algal size, odors and motility. *J Plankton Res*, 13(6):1203–1222, 1991.
Dewar WK, Bingham RJ, Iverson RL, Nowacek DP, St Laurent LC, and Wiebe PH. Does the marine biosphere mix the ocean. *J Mar Res*, 64:541–561, 2006.
Dill LM, Heithaus MR, and Walters CL. Behaviorally mediated indirect interactions in marine communities and their conservation implications. *Ecology*, 84:1151–1157, 2003.
Doall MH, Strickler JR, Fields DM, and Yen J. Mapping the free-swimming attack volume of a planktonic copepod, *Euchaeta rimana*. *Mar Biol*, 140:871–879, 2002.
Eames I, Belcher SE, and Hunt JCR. Drift, partial drift and darwins proposition. *J. Fluid Mech.*, 275:201223, 1994.
Eames I, Gobby D, and Dalziel SB. Fluid displacement by stokes flow past a spherical droplet. *J. Fluid Mech.*, 485:67–85, 2003.
Edwards AM, Phillips RA, Watkins NW, Freeman MP, Murphy EJ, Afanasyev V, Buldyrev SV, da Luz MGE, Raposo EP, Stanley HE, and Viswanathan GM. Revisiting lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449:1044 – 1048, 2007.
Euler L. Recherches générales sur la mortalité et la multiplication. *Mém. Acad. R. Sci. Bell. Lett.*, 16:144–164, 1760.
Evans GT. The encounter speed of moving predator and prey. *J Plankton Res*, 11:415–417, 1989.
Fenchel T. How dinoflagellates swim. *Protist*, 152:329–338, 2001.
Fenchel T and Blackburn N. Motile chemosensory behaviour of phagotrophic protists: mechanisms for and efficiency in congregating at food patches. *Protist*, 150:325–336, 1999.
Fenchel T and Jonsson PR. The functional biology of *Strombidium sulcatum*, a marine oligotrich

ciliate (ciliophora, oligotrichina). *Mar Ecol Prog Ser*, 48:1–15, 1988.
Fields DM and Yen J. The escape behavior of *Pleuromamma xiphias* in response to a quantifiable fluid mechanical disturbance. *Ecol Physiol*, 1:323–340, 1996a.
Fields DM and Yen J. The escape behavior of pleuromamma xiphias in response to a quantifiable fluid mechanical disturbance. *Ecol Physiol*, 1:323–340, 1996b.
Fields DM and Yen J. The escape behaviour of marine copepods in response to a quantifiable fluid mechanical distrurbance. *J Plankton Res*, 19(9):1289–1304, 1997.
Fiksen Ø, Jørgensen C, Kristiansen T, Vikebø F, and Huse G. Linking behavioural ecology and oceanography: larval behaviour determines growth, mortality and dispersal. *Mar Ecol Prog Ser*, 347:195–205, 2007.
Fisher RA. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford, UK, 1930.
Flierl GR, Grünbaum D, Levin SA, and Olson D. From individuals to aggregations: the interplay between behavior and physics. *J Theor Biol*, 196:397–454, 1999.
Folt CL and Burns CW. Biological drivers of zooplankton patchiness. *Trend Ecol Evol*, 14:300–305, 1999.
Franks PJS and Marra J. A simple new formulation for phytoplankton photoresponse and an application in a wind-driven mixed-layer model. *Mar Ecol Prog Ser*, 111:143–153, 1994.
Fung JCH. Gravitational settling of particles and bubbles in homogeneous turbulence. *J Geophys Res*, 98(C11):20287–20297, 1993.
Fung JCH. The effect of nonlinear drag on the settling velocity of particles in homogeneous isotropic turbulence. *J Geophys Res*, 103(C12):27905–27917, 1998.
Fung JCH and Vassilicos JC. Two-particle dispersion in turbulent like flows. *Phys Rev E*, 57(2): 1677–1690, 1998.
Fung JCH, Hunt JCR, Malik NA, and Perkins RJ. Kinematic simulation of homogeneous turbulence by unsteady random fourier modes. *J Fluid Mech*, 236:281–318, 1992.
Gardiner CW. *Handbook of Stochastic Methods*. Springer, Berlin, 2004.
Gargett AE. Ocean turbulence. *Annu Rev Fluid Mech*, 21:419–451, 1989.
Gargett AE, Osborn TR, and Nasmyth PW. Local isotropy and the decay of turbulence in a stratified fluid. *J Fluid Mech*, 144:231–280, 1984.
Gatignol R. The faxén formulae for a rigid sphere in an unsteady non-uniform stokes flow. *J. Méc. Théor. Appl.*, 1:143 – 154, 1983.
Gerritsen J and Strickler JR. Encounter probabilities and community structure in zooplankton: a mathematical model. *J Fish Res Board Can*, 34:73–82, 1977.
Gilliam JF. Hunting by the hunted: optimal prey selection by foragers under predation hazard. In Huges R, editor, *Behavioural Mechanisms of Food Selection*, pages 797–818. Springer-Verlag, Berlin, 1990.
Gilliam JF and Fraser DF. Habitat selection under predation hazard: test of a model with stream-dwelling minnows. *Ecology*, 68:1856–1862, 1987.
Giske J, Aksnes DL, and Førland B. Variable generation times and darwinian fitness measures. *Evolutionary Ecology*, 7:233–239, 1993.
Giske J, Rosland R, Berntsen J, and Fiksen Ø. Ideal free distribution of copepods under predation risk. *Ecol Model*, 95:45–59, 1997.
Goldstein S. On diffusion by discontinuous movements, and on the telegraph equation. *Quart J Mech and Appl Math*, 4:129–155, 1963.
Gray J and Hancock G. The propulsion of sea-urchin spermatozoa. *J Exp Biol*, 32:802–814, 1955.
Gray RD. *Foraging behavior.*, chapter Faith and foraging: a critique of the 'paradigm argument from design', page 69140. Plenum Press, New York, 1987.
Green G. Researches on the vibration of pendulums in fluid media. *Trans. R. Soc. Edinbrgh*, 13: 54–68, 1833.
Greene CH. Selective predation in freshwater zooplankton communities. *Int Rev Ges Hydrobiol*, 68: 297–315., 1983.
Grünbaum D. Advection-diffusion equations for generalized tactic searching behaviors. *J Math Biol*, 38:169–194, 1999.
Hancock G. The self-propulsion of microscopic organisms through fluids. *Proc R Soc Lond A*, 217: 96–121, 1953.

Hansen B, Bjornsen PK, and Hansen PJ. The size ratio between planktonic predators and their prey. *Limnol Oceanogr*, 39:385–403, 1994.

Hansen JLS, Kirboe T, and Alldredge AL. Marine snow derived from abandoned larvacean houses: sinking rates, particle content and mechanisms of aggregate formation. *Mar Ecol Prog Ser*, 141: 205–215, 1996.

Harris S. Steady absorption of brownian particles by a sphere. *J Chem Phys*, 77(2):934–938, 1982.

Haury LR, Kenyon DE, and Brooks JR. Experimental evaluation of the avoidance reaction of *Calanus finmarchicus*. *J Plankton Res*, 2:187–202, 1980.

Higdon JJL. A hydrodynamic analysis of flagellar propulsion. *J Fluid Mech*, 90:685–711, 1979.

Holling CS. Some characteristics of simple types of predation and parasitism. *Can Entomol*, 91: 385–398, 1959.

Holt RD. Community modules. In Begon M, Gange A, and Brown V, editors, *Multitrophic interactions in terrestrial systems*, number 17, pages 333–350. Chapman & Hall, London, 1996.

Holt RD and Polis GA. A theoretical framework for intraguild predation. *Am Nat*, 149(4):745–764, 1997.

Houston AI, McNamara JM, and Hutchinson JMC. General results concerning the trade-off between gaining energy and avoiding predation. *Phil Trans R Soc Lond B*, 341:375–397, 1993.

Hunter JR, Craig PD, and Phillips HE. On the use of random walk models with spatially variable diffusivity. *J Comput Phys*, 106:366–376, 1993.

Huntley ME and Zhou M. Influence of animals on turbulence in the sea. *Mar Ecol Prog Ser*, 273: 65–79, 2004.

Incze LS, Herbert D, Wolff N, Oakey NS, and Dye D. Changes in copepod distributions associated with increased turbulence from wind stress. *Mar Ecol Prog Ser*, 213:215–227, 2001.

Irigoien X, Harris RP, and Head RN. Does turbulence play a role in feeding and reproduction of *Calanus finmarchicus*. *J Plankton Res*, 22:399–407, 2000.

Jackson GA and Kiørboe T. Zooplankton use of chemodetection to find and eat particles. *Mar Ecol Prog Ser*, 269:153–162, 2003.

Jiang H and Osborn TR. Hydrodynamics of copepods: a review. *Surveys Geophysics*, 25:339370, 2004.

Jiang H, Meneveau C, and Osborn TR. Numerical study of the feeding current around a copepod. *J. Plankton Res.*, 21:13911421, 1999.

Jiang H, Osborn TR, and Meneveau C. Hydrodynamic interaction between two copepods: A numerical study. *J Plankton Res*, 24(3):135–253, 2002.

Jonsson PR and Tiselius P. Feeding behaviour, prey detection and capture efficiency of the copepod *Acartia tonsa* feeding on planktonic ciliates. *Mar Ecol Prog Ser*, 60:35–44, 1990.

Kamykowski D, Reed RE, and Kirkpatrick GJ. Comparison of sinking velocity, swimming velocity, rotation and path characteristics among six marine dinoflagellates. *Mar Biol*, 113:319–328, 1992.

Kamykowski D, Yamazaki H, and Janowitz GS. A lagrangian model of phytoplankton photosynthetic response in the upper mixed layer. *J Plankton Res*, 16:1059–1069, 1994.

Katija K and Dabiri JO. A viscosity-enhanced mechanism for biogenic ocean mixing. *Nature*, 460: 624–626, 2009.

Kesidis G, Konstantopoulos T, and Phoda S. Surveillance coverage of sensor networks under a random mobility strategy. *Proc IEEE Sensors*, 2:961–965, 2003.

Kessler JO. The external dynamics of swimming micro-orgainisms. *Prog Phycolog Res*, 4:258–307, 1986.

Kierstead H and Slobodkin LB. The size of water masses containing plankton blooms. *J Mar Res*, 12 (141):147–, 1953.

Kim S and Karrila S. *Microhydrodynamics.* Butterworth-Heinemann, Boston, MA, 1991.

Kiørboe T. Colonization of marine snow aggregates by invertebrate zooplankton: Abundance, scaling, and possible role. *Limnol Oceanogr*, 45(2):479–484, 2000.

Kiørboe T. Formation and fate of marine snow: small-scale processes with large-scale implications. *Sci Mar*, 65:57–71, 2001.

Kiørboe T. Sex, sex-ratios, and the dynamics of pelagic copepod populations. *Oecologia*, 148:40–50, 2006.

Kiørboe T and Saiz E. Planktivorous feeding in calm and turbulent environments, with emphasis on

copepods. *Mar Ecol Prog Ser*, 122:135–145, 1995.
Kiørboe T and Thygesen UH. Fluid motion and solute distribution around sinking aggregates. ii. implication for remote detection by colonizing zooplankters. *Mar Ecol Prog Ser*, 211:1525, 2001.
Kiørboe T and Visser AW. Predator and prey perception in copepods due to hydromechanical signals. *Mar Ecol Prog Ser*, 179:81–95, 1999.
Kiørboe T, Saiz E, and Viitasalo M. Prey switching behaviour in the planktonic copepod *Acartia tonsa*. *Marine Ecology-Progress Series*, 143(1-3):65–75, 1996.
Kiørboe T, Saiz E, and Visser AW. Hydrodynamic signal perception in the copepod *Acartia tonsa*. *Mar Ecol Prog Ser*, 179:97–111, 1999.
Kiørboe T, Grossart HP, Ploug H, and Tang K. Mechanisms and rates of bacterial colonization of sinking aggregates. *Appl Environ Microbiol*, 68:3996–4006, 2002.
Kiørboe T, Grossart HP, Ploug H, Tang K, and Auer B. Particle-associated flagellates: swimming patterns, colonization rates, and grazing on attached bacteria. *Aquat Microb Ecol*, 35:141–152, 2004.
Koehl MAR and Strickler JR. Copepod feeding currents: food capture at low reynolds number. *Limnol Oceanogr*, 26:1062–1073, 1981.
Kolmogorov AN. Local structure of turbulence in an incompressible viscous fluid at very high reynolds numbers. *Dokl. Akad. Nauk SSSR*, 301-305:30, 1941.
Koski M, Kirboe T, and Takahashi K. Benthic life in the pelagial: aggregate encounter and degradation rates by pelagic harpacticoid copepods. *Limnol Oceanogr*, 50:1254–1263, 2005.
Kozlowski J. Energetic definition of fitness? yes, but not that one. *Am Nat*, 147:1087–1091, 1996.
Krivan V and Diehl S. Adaptive omnivory and species coexistence in tri-trophic food webs. *Theor Pop Biol*, 67:85–99, 2005.
Kunze E, Dower JF, Beveridge I, Dewey R, and Bartlett KP. Observations of biologically generated turbulence in a coastal inlet. *Science*, 313:1768–1770, 2006.
Lamb H. *Hydrodynamics*. Dover, 1932.
Lampert W. The adaptive significance of diel vertical migration of zooplankton. *Funct Ecol*, 3:21–27, 1989.
Landry MR. Detection of prey by calanus pacificus: implications of the first antennae. *Limnol Oceanogr*, 25:545–549, 1980.
Lapidus R and Levandowsky M. *Biochemistry and Physiology of Protozoa*, chapter Mathematical models of behavioral responses to sensory stimuli by protozoa, page 235260. Academic Press, New York, 1981.
Leal L. *Laminar flow and convective transport processes*. Butterworth-Heinemann, Boston, MA, 1992.
Leising AW. Copepod foraging in patchy habitats and thin layers using a 2-d individual-based model. *Mar Ecol Prog Ser*, 216:167–179, 2001.
Lesieur M. *Turbulence in Fluids*. Kluwer Academic Press, 1997.
Levandowsky M, Klafter J, and White BS. Feeding and swimming behavior in grazing microzooplankton. *J Protozool*, 35(2):243–246, 1988.
Lewis DM and Pedley TJ. Planktonic contact rates in homogeneous isotropic turbulence: theoretical predictions and kinematic simulations. *J Theor Biol*, 205:377–408, 2000.
Lewis DM and Pedley TJ. The influence of turbulence on plankton predation strategies. *J Theor Biol*, 210:347–365, 2001.
Lighthill MJ. Hydromechanics of aquatic animal propulsion. *Annu Rev Fluid Mech*, 1:413–446, 1969.
Lighthill MJ. *Mathematical Biofluid Dynamics*. Society for Industrial and Applied Mathematics, Philadelphia, 1975.
Lighthill MJ. Flagellar hydrodynamics. *SIAM Rev*, 18:161–230, 1976.
Lighthill MJ. *Waves in Fluids*. Cambridge University Press, Cambridge, 1978.
Lighthill MJ. Helical distributions of stokeslets. *J Eng Math*, 30:35–78, 1996.
Lima S and Dill LM. Behavioral decisions made under the risk of predation: a review and prospectus. *Can J Zool*, 68:619–640, 1990.
Lorentz HA. Eene algemeene stelling omtrent de beweging eener vloeistof met wrijring en eenige daaruit afgeleide gevolgen. *Zittingsverskg van de Koninklijke Akademie van Wetenschappen te Amsterdam*, 5:168175, 1896.
Maar M, Nielsen TG, Stips A, and Visser AW. Microscale distribution of zooplankton in relation to

turbulent diffusion. *Limnol Oceanogr*, 48(3):1312–1325, 2003.
Maar M, Nielsen TG, Gooding S, Tönnesson K, Tiselius P, Zervoudaki S, Christou E, Sell A, and Richardson K. Trophodynamic function of copepods, appendicularians and protozooplankton in the late summer zooplankton community in the skagerrak. *Mar Biol*, 144:917–933, 2004.
Maar M, Visser AW, Nielsen TG, Stips A, and Saito H. Turbulence and feeding behaviour affect the vertical distribution of *oithona similis* and *microsetella norwegica*. *Mar Ecol Prog Ser*, 313:157–172, 2006.
Machemer H. The swimming cell and its world: structure and mechanisms of orientation in protists. *Europ J Protistol*, 37:3–14, 2001.
Mackas DL, Denman K, and Abbott MR. Plankton patchiness: Biology in the physical vernacular. *Bull Mar Sci*, 32:652–674, 1985.
MacKenzie B. Turbulence, larval fish ecology and fisheries recruitment: a review of field studies. *Oceanol Acta*, 23(4):357–375, 2000.
MacKenzie B and Kiørboe T. Larval fish feeding and turbulence: A case for the downside. *Limnol Oceanogr*, 45:1–10, 2000.
MacKenzie B, Miller T, Cyr S, and Leggett W. Evidence for a dome-shaped relationship between turbulence and larval fish ingestion rates. *Limnol Oceanogr*, 39(8):1790–1799, 1994.
Magar V, Goto T, and Pedley TJ. Nutrient uptake by a self-propelled steady squirmer. *Quart J Mech and Appl Math*, 56(1):65–91, 2003.
Malik NA and Vassilicos JC. A lagrangian model for turbulent dispersion with turbulent-like flow structure: Comparison with direct numerical simulation for two-particle statistics. *Phys Fluids*, 11 (6):1572–1580, 1999.
Mangel M and Clark CW. *Dynamic Modeling in Behavioral Ecology*. Princeton University Press, Princeton., 1988.
Mann J, Ott S, Pcseli HL, and Trulsen J. Turbulent particle flux to a perfectly absorbing surface. *J Fluid Mech*, 534:1–21, 2005.
Mariani P, MacKenzie BR, Visser AW, and Botte V. Individual based simulations of larval fish feeding in turbulent environments. *Mar Ecol Prog Ser*, 347:155–169, 2007.
Marras C, Costello JH, Granata TC, and Strickler JR. Grazing in a turbulent environment. ii energy dissipation, encounter rates and efficacy of feeding currents in centropages hamatus. *Proc Nat Acad Sci USA*, 87:1653–1657, 1990.
Maxey MR and Riley JJ. Equation of motion for a small rigid sphere in a nonuniform flow. *Phys Fluids*, 26(4):883–889, 1983.
Maxworthy T. The structure and stability of vortex rings. *J Fluid Mech*, 51:15–32, 1972.
McKendrick AG. Application of mathematics to medical problems. *Edinburgh Math. Soc.*, 44:98130, 1926.
Menden-Deuer S and Grünbaum D. Individual foraging behaviors and population distributions of a planktonic predator aggregating to phytoplankton thin layers. *Limnol Oceanogr*, 51:109–116, 2006.
Metzler R and Klafter J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339:1–77, 2000.
Michaelides EE. Review - the transient equation of motion for particles, bubbles, and droplets. *J Fluid Eng*, 119:233–247, 1997.
Mitchell JG, Pearson L, Bonazinga A, Dillon S, Khouri H, and Paxinos R. Long lag times and high velocities in the motility of natural assemblages of marine bacteria. *Appl Environ Microbiol*, 61: 877–882, 1995.
Monin P and Yaglom A. *Statistical Fluid Mechanics*. MIT Press, Cambridge MA, USA, 1975.
Munk W and Wunsch C. Abyssal recipes ii: energetics of tidal and wind mixing. *Deep-Sea Res*, 45: 1977–2010, 1998.
Okubo A. Oceanic diffusion diagrams. *Deep-Sea Res*, 18:789–802, 1971.
Okubo A. A note on small organism diffusion around an attractive center: a mathematical model. *J Oceanogr Soc Japan*, 28:1–7, 1972.
Okubo A. *Diffusion and Ecological Problems: Mathematical Models*. Springer-Verlag, Berlin, 1980.
Okubo A. Dynamical aspects of animal grouping: swarms, schools, flocks and herds. *Adv Biophys*, 22:1–94, 1986.
Okubo A and Levin SA. *Diffusion and Ecological Problems: Modern Perspectives*. Springer-Verlag,

New York, 2001.

Okubo A and Ozmidov RV. Empirical dependence of the coefficient of horizontal turbulent diffusion in the ocean on the scale of the phenomenon in question. *Oceanology*, 6:308–309, 1970.

Okuyama T and Ruyle RL. Analysis of adaptive foraging in an intraguild predation system. *Web Ecol*, 4:1–6, 2003.

Ollason JG and Lamb AE. The meaninglessness of foraging behaviour. *Perspect. Ethol.*, 11:279–296, 1995.

Osborn T. The role of turbulent diffusion for copepods with feeding current. *J Plankton Res*, 18: 185–195, 1996.

Oshanin G, Moreau M, and Burlatsky S. Models of chemical reactions with participation of polymers. *Adv Colloid Interface Sci*, 49:1–46, 1994.

Ozmidov RV. Energy distribution between oceanic motions of different scales. *Atmos. Ocean. Phys.*, 1:493–497, 1965.

Paffenhöfer GA, Strickler JR, and Alcaraz M. Suspensionfeeding by herbivorous calanoid copepods: a cinematographic study. *Mar Biol*, 67:193–199, 1982.

Patlak CS. Random walk with persistence and external bias. *Bull. Math. Biophys*, 15:311–338, 1953.

Pierce GJ and Ollason JG. Eight reasons why optimal foraging theory is a complete waste of time. *Oikos*, 49:111–118, 1987.

Pimm SL and Lawton JH. Number of trophic levels in ecological communities. *Nature*, 268:329–331, 1977.

Pimm SL and Lawton JH. On feeding on more than one trophic level. *Nature*, 275:542544, 1978.

Poisson SA. Memoire sur les mouvements simultanes d'un pendule et de l'air environnemant. *Mem. Acad. Sci. Paris.*, 9:521–523, 1831.

Poulet SA and Ouellet G. The role of amino acids in the chemosensory swarming and feeding of marine copepods. *J Plankton Res*, 4:341–361, 1982.

Purcell EM. Life at low reynolds number. *Am J Phys*, 45:3–11, 1977.

Richardson L. Atmospheric diffusion shown on a distance-neighbor graph. *Proc R Soc Lond A*, 110: 709–737, 1926.

Rothschild BJ and Osborn TR. Small-scale turbulence and plankton contact rates. *J Plankton Res*, 10:465–474, 1988.

Saito H and Kiorboe T. Feeding rates in the chaetognath sagitta elegans: effects of prey size, prey swimming behaviour and small-scale turbulence. *J Plankton Res*, 23(12):1385–1398, 2001.

Saiz E and Alcaraz M. Free-swimming behaviour of acartia clausi (copepoda: Calanoida) under turbulent water movement. *Mar Ecol Prog Ser*, 80:229–236, 1992.

Saiz E and Kiørboe T. Predatory and suspension feeding of the copepod *Arcartia tonsa* in turbulent environments. *Mar Ecol Prog Ser*, 122:147–158, 1995.

Saiz E, Alcaraz M, and Paffenhöfer GA. Effects of small-scale turbulence on feeding rate and gross-growth efficiency of three *Acartia* species (copepoda: Calanoida). *J Plankton Res*, 14:1085–1097, 1992.

Saiz E, Calbet A, and Broglio E. Effects of small scale turbulence on copepods: The case of *Oithona davisae*. *Limnol Oceanogr*, 48(3):1304–1331, 2003.

Sawford B. Turbulent relative dispersion. *Annu Rev Fluid Mech*, 33:289–317, 2001.

Schmitt FG and Seuront L. Multifractal random walk in copepod behavior. *Physica A*, 301:375–396, 2001.

Schnitzer MJ. Theory of continuum random walks and application to chemotaxis. *Phys Rev E*, 48: 2553–2568, 1993.

Seuront L, Yamazaki H, and Souissi S. Hydrodynamic distrurbance and zooplankton swimming behaviour. *Zool Studies*, 43(2):376–387, 2004.

Sharpe FR and Lotka AJ. A problem in age distribution. *Philos. Mag.*, 21:435438, 1911.

Sih A and Christensen B. Optimal diet theory: when does it work, and when and why does it fail? *Animal Behav.*, 61:379–390, 2001.

Sims DW and et al. Scaling laws of marine predator search behaviour. *Nature*, 451::1098–1102., 2008.

Skellam JG. Random dispersal in theoretical populations. *Biometrica*, 38:196–218, 1951.

Stokes GG. On the theories of internal friction of the fluids in motion. *Trans. Cambridge Philos. Soc.*, 8:287–319, 1845.

Stokes GG. On the effect of the internal friction of fluids on the motion of a pendulum. *Trans. Cambridge Philos. Soc.*, 9:8–106, 1851.

Strickler JR. Feeding currents in calanoid copepods: two new hypotheses. *Symp Soc Exp Biol*, 89: 459–485, 1985.

Strickler J. Calanoid copepods, feeding currents and the role of gravity. *Science*, 218:158–160, 1982.

Svensen C and Kiørboe T. Remote prey detection in *Oithona similis*: hydromechanical versus chemical cues. *J Plankton Res*, 22(6):1155–1166, 2000.

Tanabe K and Namba T. Omivory creates chaos in simple food web models. *Ecology*, 86:3411–3414, 2005.

Taylor GI. Diffusion by continuous movements. *Proc London Math Soc*, 20:196–212, 1921.

Taylor GI. Analysis of the swimming of microscopic organisms. *Proc R Soc Lond A*, 209:447–461, 1951.

Taylor GI. Analysis of long and narrow animals. *Proc R Soc Lond A*, 214:158–183, 1952a.

Taylor GI. The action of waving cylindrical tails in propelling microscopic organisms. *Proc Roy Soc A*, 211:225–239, 1952b.

Thar R and Fenchel T. True chemotaxis in oxygen gradients of the sulfur-oxidizing bacterium *Thiovulum majus*. *Appl Environ Microbiol*, 67:3299–3303, 2001.

Thomson DJ. Criteria for selection of stochastic models of particle trajectories in turbulent flows. *J Fluid Mech*, 180:529–556, 1987.

Tiselius P. Behavior of *Acartia tonsa* in patchy food environments. *Limnol Oceanogr*, 37(8):1640–1651, 1992.

Tiselius P and Jonsson PR. Foraging behaviour of six calanoid copepods: observations and hydrodynamic analysis. *Mar Ecol Prog Ser*, 66:23–33, 1990.

Tiselius P, Jonsson PR, Kaartveldt S, Olsen EM, and Jørstad T. Effects of copepod foraging behavior on predation risk: an experimental study of the predatory copepod *Parechaeta norvegica* feeding on *Acartia clausi* and *A.tonsa* (copepoda). *Limnol Oceanogr*, 42:164–170, 1997.

Titelman J. Swimming and escape behavior of copepod nauplii: implications for predator-prey interactions among copepods. *Mar Ecol Prog Ser*, 213:203–213, 2001.

Titelman J and Kiørboe T. Predator avoidance by nauplii. *Mar Ecol Prog Ser*, 247:137–149, 2003a.

Titelman J and Kiørboe T. Motility of copepod nauplii and implications for food encounter. *Mar Ecol Prog Ser*, 247:123–135, 2003b.

Tsinober A. *An informal introduction to turbulence*. Number 63. Kluwer Academic Press, Dordrecht, 2001.

Turchin P. *Quantitative Analysis of Movement*. Sinauer Press, Sunderland MA, 1998.

Uchaikin VV and Saenko VV. On the theory of classical mesodiffusion. *Theor Math Phys*, 46(2): 139–146, 2001.

van Duren LA. *Moving (in) water. Behavioural kinematics, energetics and hydrodynamics of the calanoid copepod Temora longicornis (Müller)*. PhD thesis, University of Groningen, 2000.

van Duren LA and Videler JJ. The trade-off between feeding, mate seeking and predator avoidance in copepods: behavioural responses to chemical cues. *J Plankton Res*, 18:805–818, 1996.

Vanderploeg HA, Paffenhfer GA, and Leibig JR. *Behavioural mechanisms for food selection.*, chapter Concentration-variable interactions between calanoid copepods and particles of different food quality: observations and hypotheses., pages 595–613. Springer, Heidelberg, 1990.

Vargas CA, Tönnesson K, Sell A, Maar M, Møller EF, Zervoudaki S, Giannakourou A, Christou E, Satapoomin S, Petersen JK, Nielsen T, and Tiselius P. Importance of copepods versus appendicularians in vertical carbon fluxes in a swedish fjord. *Mar Ecol Prog Ser*, 241:125–138, 2002.

Viitasalo M, Kiørboe T, Flinkman J, Pedersen LW, and Visser AW. Predation vulnerability of planktonic copepods: consequences of predator foraging strategies and prey sensory abilities. *Mar Ecol Prog Ser*, 175:129–142, 1998.

Visser AW. Using random walk models to simulate the vertical distribution of particles in a turbulent water column. *Mar Ecol Prog Ser*, 158:275–281, 1997.

Visser AW. Hydromechanical signals in the plankton. *Mar Ecol Prog Ser*, 222:1–24, 2001.

Visser AW. Biomixing of the oceans? *Science*, 316:838–839, 2007a.

Visser AW. Motility of zooplankton: fitness, foraging and predation. *J Plankton Res*, 29(5):447–461, 2007b.

Visser AW. Lagrangian modelling of plankton motion: From deceptively simple random walks to Fokker-Planck and back again. *Journal of Marine Systems*, 70:287–299, 2008.

Visser AW and Jackson GA. Characteristics of the chemical plume of particulate organic matter sinking in a turbulent water column. *Mar Ecol Prog Ser*, 283:55–71, 2004.

Visser AW and Jonsson PR. On the reorientation of non-spherical prey particles in a feeding current. *J Plankton Res*, 22:761–777, 2000.

Visser AW and Kiørboe T. Plankton motility patterns and encounter rates. *Oecologia*, 148:538–546, 2006.

Visser AW and MacKenzie BR. Turbulence induced contact rates of plankton: the question of scale. *Mar Ecol Prog Ser*, 166:307–310, 1998.

Visser AW and Stips A. Turbulence and zooplankton production: insights from provess. *J Sea Res*, 47:317–329, 2002.

Visser AW and Thygesen UH. Random motility of plankton: Diffusive and aggregative contributions. *J Plankton Res*, 25(9):1157–1168, 2003.

Visser AW, Saito H, Saiz E, and Kiørboe T. Observations of copepod feeding and vertical distribution under natural turbulent conditions in the north sea. *Mar Biol*, 138:1011–1019, 2001.

Visser AW, Mariani P, and Pigolotti S. Swimming in turbulence: zooplankton fitness in terms of foraging efficiency and predation risk. *J Plankton Res*, 31:121–133, 2009.

Viswanathan GM, Afanasyev V, Buldyrev SV, Murphy EJ, Prince PA, and Stanley HE. Lévy flight search patterns of wandering albatrosses. *Nature*, 381:413–415, 1996.

Viswanathan GM, Buldyrev SV, Havlin S, da Luz MGE, Raposo EP, and Stanley HE. Optimizing the success of random searches. *Nature*, 401:911–914, 1999.

von Foerster J. *The kinetics of cell proliferation*, chapter Some remarks on changing populations, page 382407. Grune and Stratton, New York, 1959.

Wunsch C. Moons, tides and climate. *Nature*, 495:743–744, 2000.

Yamazaki H and Kamykowski D. The vertical trajectories of motile phytoplankton in a wind-mixed water column. *Deep-Sea Res*, 38:219–241, 1991.

Yamazaki H, Osborn TR, and Squires KD. Direct numerical simulation of planktonic contact in turbulent flow. *J Plankton Res*, 13:629–643, 1991.

Yamazaki H, Mackas DL, and Denman K. *Biological-Physical Interactions in the Sea*, chapter Coupling small-scale physical processes with biology, pages 51–112. John Wiley and Sons, New York, 2002.

Yang DY, Makhnovskii YA, Sheu SY, and Lin SH. Simulation of the wiener sausage. *Phys Rev E*, 63 (3):3116–3120, 2000.

Yen J and Strickler JR. Advertisement and concealment in the plankton: what makes a copepod hydrodynamically conspicuous? *Invertebr Biol*, 115:191–205, 1996.

Yen J, Sanderson BG, Strickler JR, and Okubo A. Feeding currents and energy dissipation by euchaeta rimana, a subtropical pelagic copepod. *Limnol Oceanogr*, 36:362369, 1991.

Yen J, Lenz PH, Gassie DV, and Hartline DK. Mechanoreception in marine copepods: electrophysiological studies on the first antennae. *J Plankton Res*, 14:495–512, 1992.

Yen J, Weissburg MJ, and Doall MH. The fluid physics of signal perception by mate-tracking copepods. *Phil Trans R Soc Lond Biol Sci*, 353:787–804, 1998.

Young WR, Roberts AJ, and Stuhne G. Reproductive pair correlations and the clustering of organisms. *Nature*, 412:328–331, 2001.

Submitted publications

In addition to the preceding dissertation, the following articles were also submitted for consideration for the degree of Doctor Technices at the Technical University of Denmark. Much of the material in the dissertation is drawn from these articles.

- Visser AW, Mariani P, Pigolotti S. (2009) Swimming in turbulence: zooplankton fitness in terms of foraging efficiency and predation risk. *J. Plank. Res.* 31 121-133.
- Visser AW. (2008) Lagrangian modelling of plankton motion: From deceptively simple random walks to Fokker-Planck and back again. *J. Mar. Syst.* 70, 287-299.
- Visser AW. (2007) Biomixing of the oceans? *Science.* 316, 838-839.
- Visser AW. (2007) Motility of zooplankton: fitness, foraging and predation. *J. Plankton Res.* 29 (5) 447-461.
- Visser AW, Kiørboe T. (2006) Plankton motility patterns and encounter rates. *Oecologia.* 148: 539-546.
- Visser AW, Jackson GA. (2004) Characteristics of the chemical plume behind a sinking particle in a turbulent water column. *Mar. Ecol. Prog. Ser.* 283: 55-71.
- Visser AW, Thygesen UH. (2003) Random motility of plankton: Diffusive and aggregative contributions. *J. Plankton Res.* 25, 1151 - 1168.
- Visser AW. (2001) Hydromechanical signals in the plankton. *Mar. Ecol. Prog. Ser.* 222: 1-24.
- Visser AW, Saito H, Saiz E, Kiørboe, T. (2001) Observations of copepod feeding and vertical distribution under natural turbulent conditions in the North Sea. *Mar. Biol.* 138: 1011-1019.
- Visser AW, Jonsson PR. (2000) On the reorientation of non-spherical prey particles in a feeding current. *J. Plankton Res.* 22, 761-777.
- Visser AW, MacKenzie BR. (1998) Turbulence induced contact rates of plankton: the question of scale. *Mar. Eco. Prog. Ser.* 166: 307-310.
- Visser AW. (1997) Using random walk models to simulate the vertical distribution of particles in a turbulent water column. *Mar. Ecol. Prog. Ser.* 158, 275-281.

Index